

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**





## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification 5 :  <b>C12N 9/54, 15/57, 15/75</b>  <b>C11D 3/386 // (C12N 1/21</b>  <b>C12R 1/10) (C12N 1/20</b>  <b>C12R 1/07, 1/10)</b></p>	<b>A1</b>	<p>(11) International Publication Number: <b>WO 91/02792</b></p> <p>(43) International Publication Date: <b>7 March 1991 (07.03.91)</b></p>																				
<p>(21) International Application Number: <b>PCT/US90/04673</b></p> <p>(22) International Filing Date: <b>17 August 1990 (17.08.90)</b></p> <p>(30) Priority data:  <b>398,854</b>                      <b>25 August 1989 (25.08.89)</b>                      <b>US</b></p> <p>(71) Applicant: <b>HENKEL RESEARCH CORPORATION</b>  <b>[US/US]; 2330 Circadian Way, Santa Rosa, CA 95407</b>  <b>(US).</b></p> <p>(72) Inventors: <b>WILSON, Charles, R. ; 2323 Pacheco Place,</b>  <b>Santa Rosa, CA 95401 (US). LADIN, Beth, F. ; 4836</b>  <b>Fernglen Drive, Santa Rosa, CA 95405 (US). MIE-</b>  <b>LENZ, Jonathan, R. ; 1078 Ferguson Road, Sebastopol,</b>  <b>CA 95472 (US). HOM, Sherman, Siu, Ming ; 1237 St.</b>  <b>Francis Road, Santa Rosa, CA 95409 (US). HANSEN,</b>  <b>Dieter ; 800 Mission Boulevard, Santa Rosa, CA 95409</b>  <b>(US). REYNOLDS, Robert, B. ; 412 Corlano Avenue,</b>  <b>Santa Rosa, CA 95404 (US). KENNEDY, Nicholas,</b>  <b>Charles, Thomas ; Seidenweg 54, D-4000 Düsseldorf 13</b>  <b>(DE). SCHNEIDER, Joachim ; Am Eickel Kamp 158,</b>  <b>D-4010 Hilden (DE). BAHN, Michael ; Dorfstrasse 42a,</b>  <b>D-4000 Düsseldorf-Hubbeltath (DE). SCHMID, Rolf ;</b>  <b>Friedrich-Schaefer-Strasse 2, D-3340 Wolfenbüttel (DE).</b>  <b>MARKGRAF, Martina ; Kronprinzenstr. 48, D-4000</b>  <b>Düsseldorf 1 (DE). PAECH, Christian ; 2803 Audubon</b>  <b>Court, Santa Rosa, CA 95403-2460 (US).</b></p>			<p>(74) Agent: <b>DRACH, John, E.; Henkel Corporation, 300</b>  <b>Brookside Avenue, Ambler, PA 19002 (US).</b></p> <p>(81) Designated States: <b>AT (European patent), BE (European</b>  <b>patent), CA, CH (European patent), DE (European pa-</b>  <b>tent)*, DK (European patent), ES (European patent),</b>  <b>FR (European patent), GB (European patent), IT (Euro-</b>  <b>pean patent), JP, KR, LU (European patent), NL (Euro-</b>  <b>pean patent), SE (European patent).</b></p> <p><b>Published</b>  <i>With international search report.</i></p>																			
<p>(54) Title: <b>ALKALINE PROTEOLYTIC ENZYME AND METHOD OF PRODUCTION</b></p>																						
<div style="text-align: center; margin-bottom: 10px;"> <div style="display: flex; justify-content: space-around; font-weight: bold;"> <span>5</span><span>10</span><span>15</span><span>20</span><span>25</span><span>30</span> </div> <pre> 1 M K K P L G K I V A S T A L L I S V A F S S S I A S A A E E 31 A K E K Y L I G F N E Q E A V S E F V E Q V E A N D E V A I 61 L S E E E E V E I E L L H E F E T I P V L S V E L S P E D V 91 D A L E L D P A I S Y I E E D A E V T T M A Q S V P W G I S 121 R V Q A P A A H N R G L T G S G V K V A V L D T G I S T H P 151 D L N I R G G A S F V P G E P S T Q D G N G H G T H V A G T 181 I A A L N N S I G V L G V A P S A E L Y A V K V L G A D G R 211 G A I S S I A Q G L E W A G N N G M H V A N L S L G S P S P 241 S A T L E Q A V N S A T S R G V L V V A A S G N S G A S S I 271 S Y P A R Y A N A M A V G A T D Q N N N R A S F S Q Y G A G 301 L D I V A P G V N V Q S T Y P G S T Y A S L N G T S M A T P 331 H V A G A A A L V K Q K N P S W S N V Q I R N H L K N T A T 361 S L G S T N L Y G S G L V N A E A A T R </pre> </div> <div style="margin-top: 20px;"> <p>Amino acid composition</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 16.6%;">54 A</td> <td style="width: 16.6%;">0 C</td> <td style="width: 16.6%;">8 H</td> <td style="width: 16.6%;">5 M</td> <td style="width: 16.6%;">21 T</td> </tr> <tr> <td>9 R</td> <td>12 Q</td> <td>19 I</td> <td>6 F</td> <td>3 W</td> </tr> <tr> <td>24 N</td> <td>28 E</td> <td>30 L</td> <td>17 P</td> <td>9 Y</td> </tr> <tr> <td>11 D</td> <td>36 G</td> <td>10 K</td> <td>43 S</td> <td>36 V</td> </tr> </table> </div> <div style="margin-top: 10px;"> <p>Number of residues: 380  Molecular weight (MW): 38951</p> </div>			54 A	0 C	8 H	5 M	21 T	9 R	12 Q	19 I	6 F	3 W	24 N	28 E	30 L	17 P	9 Y	11 D	36 G	10 K	43 S	36 V
54 A	0 C	8 H	5 M	21 T																		
9 R	12 Q	19 I	6 F	3 W																		
24 N	28 E	30 L	17 P	9 Y																		
11 D	36 G	10 K	43 S	36 V																		
<p>(57) Abstract</p> <p>A proteolytic enzyme for use in detergent formulations having increased pH and oxidative stability under typical laundering conditions in aqueous solutions is produced by fermenting <i>Bacillus licheniformis</i> host strain transformed by a multicopy plasmid comprised of DNA sequences which code for the desired proteolytic enzyme.</p>																						

### DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

#### *FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Monaco
AU	Australia	FI	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	IT	Italy	PL	Poland
CA	Canada	JP	Japan	RO	Romania
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	LI	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
DE	Germany	LU	Luxembourg	TD	Chad
DK	Denmark			TG	Togo
				US	United States of America

## ALKALINE PROTEOLYTIC ENZYME AND METHOD OF PRODUCTION

BACKGROUND OF THE INVENTION1. Field of the Invention

This invention relates to an alkaline proteolytic enzyme, its production by cultivating genetically modified Bacillus strains, and the genetic constructs within the  
5 Bacillus strains which code for the alkaline proteolytic enzyme.

2. Description of the Related Art

Proteolytic enzymes are widely used in household laundry detergent formulations to remove protein-based  
10 stains from articles of clothing and the like. These enzymes are members of the class known as serine proteases, enzymes which catalyze the cleavage of peptide bonds. To be useful in aqueous detergent solutions, these enzymes must be active at pH values ranging from 7 to 10. They  
15 must also be stable to oxidation and enzymatic and non-enzymatic hydrolysis in order to be useful in detergent applications. Various Bacillus strains produce serine proteases under suitable aerobic fermentation conditions. However, the amount of protease made by the wild-type  
20 strains and the rate of production of the protease is

usually insufficient for commercial production. One method of increasing the protease yield is to generate mutant strains through exposure to radiation or chemical mutagens. However, these classical mutation processes suffer from the  
5 disadvantage that they produce only a statistical distribution of the desired mutation and/or that a significant population reverts back to the wild type. In recent years, efforts to overcome these problems have centered on directed genetic modification of bacterial  
10 strains. This is accomplished through the introduction of recombinant DNA into host cells under circumstances guaranteeing that the host cells will replicate the recombinant DNA with an increased copy number. In addition, the genes carried by the foreign DNA must be able  
15 to be expressed. The genetically engineered organisms do not suffer from the disadvantages often exhibited by classically mutagenized bacteria and are, therefore, often useful for commercial enzyme production.

Methods of making proteases using recombinant DNA  
20 technology are known. For example, European Patent EP130756A discloses a method for preparing carbonyl hydrolase (a protease), U.S. patent No. 4,760,025 discloses a subtilisin from having an altered amino acid sequence than that produced by the naturally occurring strain.  
25 German patent No. DE 2925427 discloses a Bacillus alkaline protease which is highly stable to chelating agents and surfactants, and is produced by fermentation of a Bacillus licheniformis strain. Production of an alkaline proteolytic enzyme by cultivation of B. licheniformis has been  
30 described in British patent No. 1,263,765; German patent No. 1,940,488; Russian patent No. 400,116; U.S. patent Nos. 3,623,956; 3,623,957.

#### SUMMARY OF THE INVENTION

It is an object of the present invention to provide a  
35 proteolytic enzyme for use in detergent formulations having increased protease and oxidative stability under conditions of pH 7-10 and at temperatures of 10-60°C in aqueous

solutions. It is also an object of the present invention to provide mutant Bacillus strains for use in the commercial production of proteolytic enzymes useful in detergent formulations. It is a further object of the present invention to provide expression vectors containing sequences which encode the desired proteolytic enzyme for transforming host Bacillus cells. These objectives are accomplished by providing a mature proteolytic enzyme having the amino acid sequence substantially as shown in Figure 1. The mature proteolytic enzyme has the amino acid sequence which extends from about amino acid residue number 112 to about amino acid residue number 380. This proteolytic enzyme is obtained by cultivating a B. licheniformis ATCC 53926 strain which has been transformed by a plasmid having a DNA sequence comprising in the direction of transcription a regulatory region, said regulatory region comprising in the order of transcription a B. licheniformis ATCC 53926 alkaline protease gene promoter, ribosome binding site, and initiation codon operably linked to a pre-pro region, said pre-pro region operably linked to a mature region encoding for the Bacillus lentus alkaline protease enzyme.

The strain Bacillus licheniformis ATCC 53926 will hereinafter be referred to as ATCC 53926. The strain Bacillus licheniformis DSM 641 will hereinafter be referred to as DSM 641. The Polymerase Chain Reaction which is described in U.S. patent Nos. 4,683,195 and 4,683,202 will hereinafter be referred to as PCR. The Bacillus licheniformis ATCC 53926 alkaline protease gene, shown in Figure 27, will hereinafter be referred to as ATCC 53926 alkaline protease gene. Bacillus lentus alkaline protease will hereinafter be referred to as BLAP. The Bacillus licheniformis ATCC 53926 alkaline protease gene promoter is the promoter from the ATCC 53926 alkaline protease gene and will hereinafter be referred to as the ATCC 53926 alkaline protease gene promoter. The abbreviation kb will hereinafter be used in place of 1000 base pairs.

BRIEF DESCRIPTION OF THE FIGURES

- Figure 1 shows the amino acid sequence for the pre, pro, and mature regions of the BLAP protein.
- Figure 2 shows the DNA sequence for the 3.4 kb HindIII fragment from Bacillus lentus which includes the BLAP gene.
- Figure 3 shows the restriction map of plasmid pGM15.
- Figure 4 shows the restriction map for plasmid pCB11C.
- Figure 5 shows the restriction map for plasmid pCB1N.
- Figure 6 shows the restriction map for plasmid pCB13C.
- Figure 7 shows the restriction map for plasmid pCB2N.
- Figure 8 shows the oligonucleotides 1-3. Positions correspond to the positions of the nucleotide sequences within the ATCC 53926 alkaline protease gene. Mismatch positions are indicated by the use of lower case letters.
- Restriction sites are underlined and indicated.
- Figure 9 shows the restriction map for plasmid pCB15C.
- Figure 10 shows the restriction map for plasmid pCB4N.
- Figure 11 shows a scheme for the construction of plasmid pCB1N.
- Figure 12 shows a scheme for the construction of plasmid pCB11C.
- Figure 13 shows the restriction map for plasmid pCB55C.
- Figure 14 shows the restriction map for plasmid pCB56C.
- Figure 15 shows the restriction map for plasmid pCB58C.
- Figure 16 shows the restriction map for plasmid pCB75C.
- Figure 17 shows the restriction map for plasmid pCB78C.
- Figure 18 shows the restriction map for plasmid pCB50N.
- Figure 19 shows the restriction map for plasmid pCB51N.
- Figure 20 shows the restriction map for plasmid pCB53N.
- Figure 21 shows the restriction map for plasmid pCB70N.
- Figure 22 shows the restriction map for plasmid pCB73N.
- Figure 23 shows the scheme for construction of the ClaI fusion plasmid constructs in Escherichia coli and Bacillus.
- Figure 24 shows a schematic diagram of the Savinase oligonucleotide probes. These probes were designed on the basis of known Savinase amino acid sequence to probe the



Bacillus lentus gene banks and to serve as primers in the PCR reaction.

Figure 25 shows the specific sequence of oligonucleotide probes and primers.

- 5 Figure 26 shows the restriction map for plasmid pMG108.2  
Figures 27 and 27a show the introduction of restriction sites into the gene encoding the ATCC 53926 alkaline protease gene.

- Figure 28 shows a comparison of the relative protease  
10 yields of the different ATCC 53926 alkaline protease-BLAP constructs in the ATCC 53926 production strain.

Figure 29 shows the amino acid sequence of the mature protease made from the ATCC 53926 alkaline protease-BLAP NcoI gene fusion constructs.

15 DESCRIPTION OF THE PREFERRED EMBODIMENTS

- A proteolytic enzyme (BLAP) produced by B. lentus, is provided having increased protease and oxidative stability in aqueous solutions of pH 7-10 and at temperatures of from 10-60°C. The mature BLAP enzyme is a polypeptide having  
20 269 amino acid residues, a molecular weight of 26,823 daltons, and a calculated isoelectric point of 9.7. It was first obtained from B. lentus DSM 5483, an organism isolated from soil samples and having the properties listed in Table I.

25 Table I. Properties of B. lentus DSM 5483

TAXONOMIC NAME	<u>B. lentus</u> (Group II)
MORPHOLOGY	Vegetative Cells:
	Rods
	Width 0.6-0.8 $\mu\text{m}$
30	Length 1.5-3.5 $\mu\text{m}$
	Spores:
	Ellipsoid
	Sporangium not swollen
35	
MOBILITY	Motile
GRAM REACTION	Positive
40	
CATALASE	Positive
ANAEROBIC GROWTH	Negative

	VOGUES-PROSKAUER REACTION	Negative pH in Vogues-Proskauer medium: 6.2	
5	GROWTH PARAMETERS	TEMPERATURE	
		Growth	50°C
		No growth	55°C
		Optimal	39°C
10		pH:	
		Growth	pH 7.5 - pH 9.5
		No growth	pH 5.7
		SALT TOLERANCE	
15		Grows in	5% NaCl
		Grows in	7% NaCl
		Grows in	10% NaCl
		SUITABLE MEDIUM:	
20		Potato Starch	10 g/l
		Yeast Extract	10 g/l
		Peptone	5 g/l
		Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> ·5H <sub>2</sub> O	2.5 g/l
		K <sub>2</sub> HPO <sub>4</sub>	2 g/l
25		Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O	1.2 g/l
		NH <sub>4</sub> Cl	1 g/l
		FeCl <sub>3</sub> ·6H <sub>2</sub> O	540 mg/l
		MgO	500 mg/l
		CaCO <sub>3</sub>	200 mg/l
30		MnCl <sub>2</sub> ·4H <sub>2</sub> O	100 mg/l
		ZnO	40 mg/l
		CoCl <sub>2</sub> ·6H <sub>2</sub> O	24 mg/l
		CuCl <sub>2</sub> ·2H <sub>2</sub> O	20 mg/l
35		H <sub>3</sub> BO <sub>4</sub>	3 mg/l
40	ACID PRODUCTION FROM:	Glucose	(weak)
		L-Arabanose	(weak)
		Xylose	(weak)
		Mannitol	(weak)
	GAS PRODUCTION FROM GLUCOSE	Negative	
45	LECITHINASE	Negative	
	HYDROLYSIS	Starch:	Positive
		Gelatin:	Positive
		Casein:	Positive
50	ACID UTILIZATION	Citrate:	Positive
		Propionate:	Negative
	TYROSINE DEGRADATION	Negative	
55	INDOLE REACTION	Negative	

	PHENYLALANINE DEAMINASE	Negative
	ARGININE DEHYDROLASE	Negative
5	NITRATE REDUCTION	Nitrate reduced to nitrite

The amino acid sequence, shown in Figure 1 was determined in part by standard amino acid sequencing techniques, and the rest deduced from the nucleic acid sequence.

10 A method is presented for producing the BLAP enzyme comprising cultivating a Bacillus host cell containing a hybrid plasmid having a DNA sequence coding for the BLAP enzyme comprising in the direction of transcription, a  
15 regulatory region, said regulatory region comprising in the order of transcription a ATCC 53926 alkaline protease gene promoter, ribosome binding site, and initiation codon operably linked to a pre-pro region, said pre-pro region operably linked to a mature region encoding for the BLAP  
20 enzyme. All the above may be linked to a carboxy terminal (C-terminal) DNA sequence from ATCC 53926 alkaline protease gene. The C-terminal DNA sequence is a DNA sequence between approximately the SalI and approximately the SstI restriction sites of the ATCC 53926 alkaline protease gene.  
25 It may be used to suppress chromosomally encoded protease as disclosed in a copending U.S. patent application. Any Bacillus strain can be used in the method of the present invention to produce the BLAP enzyme. B. licheniformis is the preferred strain. ATCC 53926 is the most preferred  
30 strain. ATCC 53926 is a mutant strain obtained by classical mutagenesis of B. licheniformis DSM 641 having the properties listed in Table II.

Table II. Properties of B. licheniformis ATCC 53926

	TAXONOMIC NAME	<u>B. licheniformis</u> ATCC 53926	
	MORPHOLOGY	Vegetative Cells:	
		Rods	
5		Width	0.7-0.9 $\mu\text{m}$
		Length	2.0-3.0 $\mu\text{m}$
		Spores:	
		Ellipsoid	
10		Sporangium not swollen	
	MOBILITY	Motile	
	GRAM REACTION	Positive	
15			
	CATALASE	Positive	
	ANAEROBIC GROWTH	Positive	
20	VOGUES-PROSKAUER REACTION	Positive	
		pH in Vogues-Proskauer medium: 5.5	
	GROWTH PARAMETERS		
		TEMPERATURE	
25		Growth	50°C
		No growth	55°C
		pH:	
		No growth	pH 5.7
30		SALT TOLERANCE	
		Grows in 5% NaCl	
		Grows in 7% NaCl	
		Grows in 10% NaCl	
35	ACID PRODUCTION FROM:	Glucose	positive
		L-Arabanose	positive
		Xylose	positive
		Mannitol	positive
40	GAS PRODUCTION FROM GLUCOSE	Negative	
	LECITHINASE	Negative	
	HYDROLYSIS	Starch:	Positive
45		Gelatin:	Positive
		Casein:	Positive
	ACID UTILIZATION	Citrate :	Positive
		Propionate:	Negative
50			
	TYROSINE DEGRADATION	Negative	
	INDOLE REACTION	Negative	
55	PHENYLALANINE DEAMINASE	Negative	

ARGININE DEHYDROLASE	Negative
NITRATE REDUCTION	Nitrate not reduced to nitrite

5       The gene coding for the BLAP enzyme can be obtained by isolating the chromosomal DNA from the B. lentus strain DSM 5483, constructing DNA probes having homology to putative DNA sequences encoding regions of the B. lentus protease, preparing genomic libraries from the isolated chromosomal  
10 DNA, and screening the libraries for the gene of interest by hybridization to the probes. The genomic libraries can be constructed in E. coli using plasmid pBR322, bacteriophage lambda EMBL3 and EMBL4, and the cosmid pCP13 as cloning vectors. Preferably, the gene bank of B. lentus  
15 DSM 5483 chromosomal DNA is constructed in E. coli with the cosmid pCP13.

Oligonucleotide probes having DNA sequences similar to those of the Savinase gene (a gene encoding for Savinase, a commercial serine protease sold by Novo Industri A/S) can  
20 be used to probe for Savinase-like genes. Preferably, a BLAP sequence can be used. These oligonucleotides can be used to screen Southern blots of restriction enzyme digested chromosomal DNA's from E. coli HB101, B. subtilis DB104, B. licheniformis, and the B. lentus strain. The N-  
25 terminal and internal oligonucleotides as shown in Figures 24 and 25 can also be utilized as primers in the PCR to amplify that region of B. lentus chromosomal DNA extending from the N-terminus of the mature alkaline protease to the N-terminus of the C-terminal cyanogen bromide fragment.

30       A portion of the mature region of the alkaline protease gene can be isolated by constructing a larger, double stranded DNA probe specific for the B. lentus alkaline protease using the PCR. An analysis based on the amino acid sequence for the C-terminal cyanogen bromide  
35 fragment from the Savinase protease shows that this sequence has homology with the amino acid sequence around the active site serine of Savinase, as well as other subtilisin type serine proteases. The amplified fragment

is approximately 650 basepairs (bp) in length, in agreement with the expected size estimated from sequence data available for other Bacillus alkaline proteases. The 650 bp PCR fragment represents a portion of a gene encoding an alkaline protease, beginning just after the N-terminus of the mature protease and continuing into the protein coding region.

The 650 bp fragment can be used to identify the BLAP gene in the genomic library. After purification, the 650 bp fragment can be labeled for example by nick translation and used to probe gene banks of B. lentus DSM 5483 DNA which may be constructed by using the cosmid pCP13 and EMBL lambda vectors. Several cosmid clones carrying DNA with homology to the 650 bp PCR fragment and to selective oligonucleotide probes may then be identified. Plasmid DNA of one of these cosmids, pSH44, has been purified and characterized by restriction enzyme analysis. Results from Southern analysis show that a B. lentus DSM 5483 HindIII fragment of approximately 3.4 kb can hybridize to the radioactively labeled probes 650 bp PCR fragment. The hybridization, restriction analyses and DNA sequence (Figure 2) show that the 3.4 kb HindIII fragment having homology to the 650 bp PCR fragment encodes the BLAP gene.

The fermentation of ATCC 53926 containing an expression cassette which is based on gene fusions comprising parts of the ATCC 53926 alkaline protease gene and parts of the BLAP gene is used in the preferred method of making BLAP on a large scale production basis. Such a cassette consists of a regulatory region from the ATCC 53926 alkaline protease gene comprised of a promoter, ribosome binding site and initiation codon which is responsible for the efficient transcription and translation of the corresponding gene. The regulatory region is followed by a DNA sequence coding for the signal (pre) peptide, followed by sequences encoding the pro-peptide and the mature protease. After transcription of the DNA sequence downstream from the promoter, the messenger-RNA

(m-RNA) is translated into the pre-pro-mature protease, which consists of the signal peptide, the pro peptide and the mature protease. The signal sequence directs the immature protease through the cytoplasmic membrane, where the signal peptide is cleaved off enzymatically by a signal peptidase. The pro sequence seems to insure correct folding and processing of the pro-protease during or after translocation across the cytoplasmic membrane and the cell wall, releasing the mature protease (Power, S. D. *et al.* (1986) PNAS(USA), 83, 3096; Ikemura, H. *et al.* (1987) J. Biol. Chem. 262, 7859; Ikemura, H. *et al.* (1988) J. Biol. Chem. 263, 12959).

Two fusion types identified by the restriction enzyme sites used to join different portions of the ATCC 53926 alkaline protease gene and BLAP genes together can be used to prepare the previously described expression cassettes which when used to transform ATCC 53926 will produce BLAP enzyme on a commercial scale. These cassettes are made from ClaI and NcoI fusions.

The ClaI fusion combines the ATCC 53926 alkaline protease gene promoter and pre sequence with a DNA fragment encoding the pro and mature sequence of BLAP. A ClaI restriction site can be introduced into the ATCC 53926 alkaline protease gene several codons upstream from the junction of the ATCC 53926 alkaline protease pre and pro sequence. The BLAP gene contains a naturally occurring ClaI site at this position. The promoter and most of the pre region of the ATCC 53926 alkaline protease gene (AvaI/ClaI DNA fragment) can be fused to the (ClaI/SstI) fragment of the BLAP gene which contains a portion of the pre region and all of the pro and mature BLAP sequence. The NcoI and ClaI fusions can be made in E. coli and subcloned into two different Bacillus vectors on an AvaI/SstI DNA fragment.

The NcoI fusion produces a mature protease slightly different from the BLAP. The protein made from the NcoI fusion contains threonine at position 3 from the N-terminus

of the mature protease while the BLAP contains serine at the same position (Figure 29). The NcoI fusion can be constructed by combining the promoter, pre and pro sequences of the ATCC 53926 alkaline protease gene with the  
5 mature sequence of the BLAP gene. A NcoI restriction site can be introduced into the ATCC 53926 alkaline protease gene several codons downstream from the junction of the pro and mature sequence. A naturally occurring NcoI site is present at a similar position in the BLAP gene. Then, the  
10 promoter, pre, pro and four codons of the mature ATCC 53926 sequence (AvaI/NcoI DNA fragment) can be fused to a mature BLAP sequence missing the first four codons (NcoI/SstI DNA fragment). The NcoI and ClaI fusions can be made in E. coli and subcloned into two different Bacillus vectors on an  
15 AvaI/SstI DNA fragment.

To further increase enzyme production the ClaI and NcoI fusions can be cloned either with a ATCC 53926 alkaline protease transcription terminator cassette which is a 164 bp DNA fragment, or with a 608 bp C-terminal  
20 fragment from the ATCC 53926 alkaline protease gene which is discussed in more detail below.

The ATCC 53926 alkaline protease-BLAP fusion genes can be constructed in the E. coli vector pGM15 (Figure 3), a pTZ19R derivative that contains the 3.4 kb HindIII/KpnI  
25 fragment containing the BLAP gene, and then subcloned into Bacillus vectors pBC16 and pUB110. Bacillus fusion gene vector pCB1N (Figure 5) is constructed as shown in Figure 11. Bacillus fusion gene vector pCB11C (Figure 4) can be constructed as shown in Figure 12. Variations of each of  
30 the gene fusions described above can be prepared in order to: (a) provide more effective expression of the BLAP protease gene, and (b) to ensure suppression of the expression of the chromosomally encoded ATCC 53926 alkaline protease gene. In order to provide effective termination  
35 of the transcription of the BLAP protease gene, the ATCC 53926 alkaline protease transcription terminator can be cloned as a 164 bp DNA fragment downstream of the BLAP



terminator in each of the gene fusions described above. Constructs carrying the ATCC 53926 alkaline protease transcription terminator cassette in the same orientation relative to the transcription of the ATCC 53926 alkaline protease-BLAP fusion gene in plasmids pCB11C and pCB1N are designated pCB13C and pCB2N (Figure 6 and 7) respectively. In order to suppress the expression of the chromosomally encoded ATCC 53926 protease gene, two additional Bacillus BLAP fusion gene vectors containing a C-terminal fragment of the ATCC 53926 alkaline protease gene can be constructed in plasmids pCB11C and pCB1N by ligating a 608 bp fragment of the ATCC 53926 alkaline protease gene synthesized by the PCR using plasmid pC51 as template. The 608 bp DNA fragment can be ligated into the single KpnI site present on vectors pCB11C and pCB1N and new constructs containing the 608 bp ATCC 53926 C-terminal fragment in the same orientation relative to the direction of transcription of the ATCC 53926 alkaline protease-BLAP fusion genes in pCB11C and pCB1N are designated pCB15C and pCB4N respectively. Restriction maps of these plasmids are given in Figures 9 and 10 respectively.

The six ATCC 53926 alkaline protease-BLAP fusion genes described above and listed below can be subcloned into Bacillus vectors and transformed into B. subtilis DB104. The ATCC 53926 alkaline protease-BLAP fusion genes can be subcloned as AvaI/SstI DNA fragments into pBC16 and pUB110 derivatives by digesting the E. coli vectors containing the ATCC 53926 alkaline protease-BLAP fusion genes (pCB1N, pCB11C, pCB2N, pCB13C, pCB4N and pCB15C) with AvaI and SstI, and ligating them to Bacillus plasmids pC51 and pH70 which have been restricted with AvaI, SstI, PstI and SalI. Plasmids pC51 and pH70 were cut with AvaI/SstI to provide the large AvaI/SstI vector fragment containing the necessary origin of replication and an antibiotic resistance gene (tetracycline resistance in pC51, kanamycin resistance in pH70) for selection. SalI and PstI can be used to restrict the small AvaI/SstI fragment containing

the entire ATCC 53926 alkaline protease gene present on pC51, and pH70 into three smaller DNA fragments. This decreases the probability of religating the ATCC 53926 alkaline protease gene into the Bacillus vectors.

- 5 A total of ten ATCC 53926 alkaline protease-BLAP gene fusion plasmids constructed as described above derived from the ClaI and NcoI fusions can be realized in ATCC 53926. Five Bacillus plasmid constructs derived from the six ClaI fusions can be transformed into the ATCC 53926 strain.
- 10 These Bacillus plasmids are pCB55C, pCB56C, pCB58C, pCB75C, and pCB78C (Figures 13, 14, 15, 16, and 17 respectively). Five Bacillus plasmid constructs derived from the six NcoI fusions can be transformed into the ATCC 53926 strain. These Bacillus plasmids are pCB50N, pCB51N, pCB53N, pCB70N,
- 15 and pCB73N (Figures 18, 19, 20, 21, and 22 respectively).

For production of the greatest protease yields, the ATCC 53926 strain carrying the ClaI fusion constructs (Figure 23) inserted into a pC51 derivative is preferred. Construct pCB58C (Figure 15) in the ATCC 53926 strain is

20 the most preferred embodiment for producing the greatest protease yields.

The fermentations can be carried out in complex nutrient media generally known to those skilled in the art. These media can contain such carbon sources as glucose,

25 sucrose, and/or corn starch. The media can also contain such nitrogen sources as soy bean meal, cottonseed meal, casein and potatoe extract. Other growth promoting ingredients include brewing residue, and corn steep liquor and inorganic materials such as inorganic phosphates. The

30 media can also contain anti-foams.

A comparison of the relative protease yields of the different ATCC 53926 alkaline protease-BLAP constructs in the ATCC 53926 production strain is shown in Figure 28. These fermentations were performed with antibiotic

35 selection, using a complex production medium containing 0.5% nitrogen. Protease yields of the ATCC 53926 alkaline protease-BLAP constructs are presented relative to ATCC

53926 containing plasmid pAM2.3. This plasmid contains the native BLAP promoter and structural gene on an NruI/SalI restriction fragment inserted into plasmid pC51 between the StuI/SalI sites.

5 The fermentation broths obtained by cultivating the different ATCC 53926 alkaline protease-BLAP constructs in the ATCC 53926 production strain typically contain other proteolytic enzymes. However, the BLAP protease in the fermentation broths obtained by fermentation of the  
10 different ATCC 53926 alkaline protease-BLAP constructs in the ATCC 53926 production strain is the major proteolytic enzyme in the broths. These fermentation broths contain proteolytic enzymes not present in broths obtained by cultivating Bacillus lentus DSM 5483. Specifically,  
15 fermentation broths obtained by fermentation of the different ATCC 53926 alkaline protease-BLAP constructs in the ATCC 53926 production strain produce the BLAP protease and at least one other proteolytic enzyme produced by the ATCC 53926 strain. Preferably, the BLAP enzyme is obtained  
20 as a composition comprising the BLAP enzyme which is a mature alkaline proteolytic enzyme from Bacillus lentus DSM 5483 having the amino acid sequence from about amino acid residue 112 to about amino acid residue 380 substantially as shown in Figure 1 and at least one other proteolytic  
25 enzyme from Bacillus licheniformis ATCC 53926.

The proteases produced by cultivating the ATCC 53926 strain containing the various recombinant plasmids are particularly useful in detergent formulations. They can be formulated with surfactants in accordance with methods  
30 known in the art of formulating detergent and cleaning compositions especially laundry detergent compositions. A surfactant is any substance which reduces the surface free energy of water including soaps, nonionic surfactants, anionic surfactants, and cationic surfactants. In the case  
35 of laundry detergent formulations, the enzymes can be combined with such suitable surfactants as linear alkyl benzene sulfonates, ethoxylated linear alcohols,

ethoxylated alkyl sulfate, or sulfated linear alcohol. The formulation can also contain builders, fluorescent brighteners, bleach, and other components normally used in the art.

5

#### DEPOSIT OF MICROORGANISMS

Living cultures of ATCC 53926 and plasmid pCB58C in ATCC 53926, assigned the ATCC Designation 68032, have been accepted for Deposit under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms  
10 for the purposes of patent procedure by the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852. A living culture of B. lentus DSM 5483 has been accepted for deposit in Deutsche Sammlung von Mikroorganismen (DSM), Grisebachstr. 8, 3400 Göttingen,  
15 Federal Republic of Germany, under the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the purposes of patent procedure.

The following examples will serve to illustrate but not to limit the invention.

20

#### Example 1.

##### Construction of a B. lentus DSM 5483 Gene Bank

##### A. Using the Cosmid Cloning Vector pCP13

The cosmid cloning vector pCP13 was chosen to construct a genomic library from B. lentus DSM 5483 DNA.  
25 This broad host range vector derived from the plasmid pLAFR1 (Friedman et. al. Gene(1982) 18, 289) was obtained from F. Ausubel, Department of Molecular Biology, Massachusetts General Hospital, Boston, MA. Plasmid pCP13 is 23 kb in size and encodes both tetracycline ( $Tc^R$ ) and  
30 kanamycin ( $Km^R$ ) resistance, contains multiple unique endonuclease restriction sites, and is mobilizable but not self-transmissible. Cloning into the BamHI site on pCP13 leads to insertional inactivation of the Km resistance determinant, allowing recombinants to be identified by  
35 their  $Tc^R, Km^S$  phenotype. The vector also contains the lambda cos sites which allows concatemeric forms of the cosmid to be packaged into lambda phage particles using

commercially available packaging systems. Inserts in the size range of 12 to 30kb can be cloned into pCP13, making this cloning vector very attractive for the construction of gene banks. Purified pCP13 DNA (20 $\mu$ g) was digested with

5 BamHI (3 units (U)/ $\mu$ g) and subsequently dephosphorylated to prevent recircularization. Purified chromosomal DNA (410  $\mu$ g) from B. lentus strain DSM 5483 was partially digested with BclI (0.0625U/ $\mu$ g DNA) at 50°C for 60 minutes, and then ligated to the BamHI digested pCP13 DNA. The ligation

10 reaction contained 150  $\mu$ g/ml pCP13 vector, and 50  $\mu$ g/ml DSM 5483 chromosomal DNA in a final volume of 30  $\mu$ l. Ligations were held for 9 hours at 14°C. The ligation mixtures containing concatemeric pCP13-B. lentus genomic DNA fragments were packaged into lambda phage heads using

15 the Packagene Lambda DNA packaging System purchased from Promega, Madison, Wisconsin. The recombinant cosmid molecules packaged into phage particles were used to infect E. coli HB101. E. coli HB101 was grown overnight at 30°C in a rotary shaker at 175rpm in 10 ml of Tryptone Broth or

20 TB (per/liter (l), 10 g tryptone, 5 g NaCl, 2 g maltose and 2.46 g Mg<sub>2</sub>SO<sub>4</sub>, sterilized by autoclaving). The cells were pelleted and resuspended in 10 mM Mg<sub>2</sub>SO<sub>4</sub>. Packaging mixture and cells (50 $\mu$ l and 100 $\mu$ l respectively) were mixed together and incubated at 37°C for 20 minutes. Then 1 ml of Luria

25 Broth (per/l, 10 g Bactotryptone, 10 g NaCl, 5 g Yeast extract) was added, and the tubes incubated at 37°C for 60 minutes in New Brunswick G24 rotary shaker at 150rpm. The infected HB101 cells were pelleted by centrifugation (13000 x g) for 5 min at room temperature, resuspended in Luria

30 Broth, and plated onto Luria agar containing 15 $\mu$ g/ml Tc. Fourteen hundred Tc<sup>R</sup> Km<sup>S</sup> HB101 recombinant clones were obtained. To determine the average insert size, plasmid DNA from 14 randomly selected potential clones was isolated, digested with EcoRI, and analyzed by agarose gel

35 electrophoresis. Eleven of the fourteen clones contained inserts with an average insert size of 25.9kb (range 21.7 to 32.7kb). Based on this large average insert size, and

the assumption that the B. lentus genome is approximately the same size as the E. coli genome, theoretical calculations predicted a 99% probability of detecting the Savinase-like gene by screening 768 cosmid clones.

5        **B. Using the Phage Cloning Vectors EMBL3 and EMBL4**

Another method of cloning large pieces of genomic DNA (>10kb), is to use bacteriophage lambda cloning vectors. These insertion vectors can accept between 9kb and 23kb of heterologous DNA. The heterologous DNA and the lambda DNA  
10 are necessary to package the recombinant DNA into phage heads which can then be propagated in an E. coli host. The EMBL3 and EMBL4 phage vectors were purchased from Stratagene in La Jolla, CA. These cloning vectors have been constructed with a polylinker region flanking the  
15 internal stuffer DNA which is not needed for replication. When the lambda DNA is restricted with the appropriate enzyme, the remaining two DNA fragments called the left and the right arms needed for replication, can be ligated to heterologous DNA. In order for the recombinant DNA to be  
20 efficiently packaged into the phage heads, the ligations must be performed at high DNA concentrations to favor the formation of concatamers, and the heterologous insert must be between 9kb and 23kb in size. This latter requirement is based on the observation that the length of the DNA  
25 which can be packaged is between 70% and 105% of the size of the wild-type lambda genome. Further, both EMBL systems take advantage of spi selection (sensitive to bacteriophage P2 inhibition). EMBL arms ligated to the stuffer fragment and packaged into phage heads cannot grow on a host strain  
30 containing a P2 lysogen. In this system, it is unlikely that stuffer fragment and EMBL arms do religate. If this occurs, these phage will not replicate when plated on the appropriate E. coli host.

In order to prepare genomic libraries of DSM 5483 DNA  
35 in either the EMBL3 or EMBL4 vectors, DSM 5483 DNA was partially digested with either EcoRI or BclI (the 5' sticky end produced by BclI is complementary to the ends of the

vector cut with BamHI). In each case, the genomic DNA was digested with less than 1 unit of enzyme per microgram of DNA for 60 minutes. Optimal packaging efficiencies are obtained with lambda DNA's that are concatemeric.

5 Therefore, the ligation reactions were set to achieve DNA concentrations of 200µg/ml, which favors concatemer formation. In a standard 5µl ligation reaction, 1µg of vector arms was mixed with 200 to 400 ng of genomic DNA. Ligation for 1 hour at room temperature was followed by an

10 overnight incubation at 4°C. Concatemeric DNA from the ligation reactions was packaged into phage heads using GigaPack Plus in vitro packaging extract purchased from Stratagene (La Jolla, California), and the protocol provided with the extracts. After packaging, 0.5ml of

15 phage dilution buffer was added to the packaging reactions (per liter, 5.0 g NaCl, 2 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 50ml 1M Tris HCl pH7.5 and 5ml 2% gelatin) and 20µl of chloroform. The packaging reactions were titered, and amplified according to the protocols provided by Stratagene. From each phage

20 library 5 X 10<sup>5</sup> recombinant phage were obtained. When a similar experiment was performed with a test insert provided with the commercial cloning kit, 3 X 10<sup>6</sup> recombinant phage were obtained. Less than 5% of the plaques generated by the control insert were due to

25 background phage. Recombinant phage had an average insert size of 12-14kb with about 70% of the phage actually showing inserts. The phage libraries were amplified to a titer of 10<sup>8</sup>phage/ml and stored over chloroform at 4°C. Other stocks were prepared for longer term storage at -

30 70°C.

To characterize the EMBL libraries and to find the Savinase-like gene, the following approach was undertaken. Groups of plaques were isolated from either the EMBL3 or the EMBL4 libraries, and their DNA was isolated by a phage

35 minipreparation procedure (Maniatis et. al., (1982), A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, New York). The insert genomic DNA present

in the EMBL3 and EMBL4 libraries was digested with either SalI or EcoRI, respectively, and after agarose gel electrophoresis the DNA fragments were transferred to Gene Screen Plus. Total DSM 5483 genomic DNA was nick translated and used as a hybridization probe to verify that the inserts present in the EMBL libraries were really DSM 5483 DNA. As expected, the cloned inserts hybridized with the labelled DSM 5483 DNA. In a second experiment, genomic DSM 5483 DNA was cut with either EcoRI or BclI, electrophoresed in an agarose gel, and then transferred to Gene Screen. DNA isolated from groups of recombinant phage from either the EMBL3 or EMBL4 libraries was nick translated and used as a hybridization probe. As expected, hybridization to DSM 5483 genomic DNA occurred while little or no hybridization was detected to DNA from other Bacillus species. These results verified the formation of gene libraries of B. lentus DSM 5483 chromosomal DNA using the lambda EMBL cloning vectors.

#### Example 2.

#### 20 Construction of DNA Probes Specific for the Savinase-like Protease Gene Present in B. lentus DSM 5483

##### A. Oligonucleotide Probes

The gene bank of DSM 5483 chromosomal DNA constructed in E. coli with the cosmid pCP13 was screened using oligonucleotide probes. A series of probes of varying degeneracy were synthesized using a Biosearch DNA synthesizer. The sequence of these oligonucleotides was deduced from amino acid sequence data obtained from the N-terminus of the Savinase protease and from the C-terminal fragment of Savinase derived by cyanogen bromide cleavage. The sequence and location of these probes in relation to the Savinase protease are depicted in Figures 24 and 25. The NO2 and NO4 series of probes were both designed from N-terminal amino acid sequence data. The NO2 series was 23 bases in length with 14 bases of homology and contained 16 fold degeneracy, while the NO4 series were 17 bases in length with 48 fold degeneracy. The NO4.1-4 series of



probes was designed using the most probable codon usage for B. subtilis. The CO3 series of oligonucleotides were all 17 bases in length and had a 32 fold degeneracy. These latter probes were designed from amino acid sequence derived from the C-terminal cyanogen bromide fragment of Savinase. These probes were purified, radioactively labeled, and hybridized to chromosomal DNA from DSM 5483 which had been restricted, electrophoresed on agarose gels, and blotted onto nitrocellulose using the procedure of Southern. The restriction enzymes used included EcoRI, HindIII, TaqI, Sau3A, SstI, XbaI, and XhoI. The strongest hybridization signals were obtained with oligonucleotides NO2.2, NO4.2, and CO3.3 (Figure 25). The latter three labeled oligonucleotides were used to screen Southern blots of restricted chromosomal DNA's from E. coli HB101, B. subtilis DB104, B. licheniformis, and the B. lentus strain, and were found to hybridize only to DNA fragments from the B. lentus strain. This result confirmed that these three oligonucleotides were specific for DNA encoding the Savinase-like protease.

#### B. Amplification of a Portion of the Alkaline Protease Gene Sequence

Although three of the oligonucleotide probes described above were specific for DSM 5483 chromosomal DNA an additional approach was pursued in order to isolate a larger, double stranded DNA probe specific for the Savinase-like alkaline protease. The N-terminal and internal oligonucleotides described above were utilized as primers in the PCR to amplify that region of DSM 5483 chromosomal DNA extending from the N-terminus to amino acid residue 222 of the mature alkaline protease (Figure 24). The amino acid sequence obtained from the C-terminal cyanogen bromide fragment from Savinase had homology with the amino acid sequence around the active site region of other serine proteases. If the distance from the N-terminus of Savinase to the active site region was similar to other serine

proteases, then the expected size of the amplified fragment would be approximately 650 bp.

The PCR was carried out using a commercial GeneAmp kit purchased from Perkin - Elmer Cetus Instruments. In the  
5 PCR, the DNA segment to be amplified may range from a few hundred bp to more than 5 kb in length, and is contained on either a plasmid or on a chromosomal DNA template. Amplification requires the synthesis of two oligonucleotide primers, approximately 15 to 25 bases in length, designed  
10 to have homology to opposite DNA strands at either end of the DNA segment. Since DNA synthesis occurs in a 5' to 3' direction, each primer is designed to mimic the corresponding 5' to 3' strand. In the initial experiment, 16 different PCR's were performed using all possible  
15 combinations of probe series CO3 with NO2, and DSM 5483 chromosomal DNA as template. The concentration of template and primers was as prescribed in the kit. The PCRs were initiated by mixing template, excess primers, dNTP's, heat stable DNA polymerase, and reaction buffer in eppendorf  
20 tubes which were cycled for set periods of time at 94°C, 37°C, and 72°C. In each cycle, the double stranded DNA template was denatured by heating at 94°C, then cooled to 37°C to allow for annealing of the primers to the appropriate DNA sequences on the single stranded template.  
25 The tubes were then shifted to 72°C, where complementary DNA strands were synthesized through elongation of the annealed primers by a heat stable DNA polymerase (Taq polymerase). This process was repeated 25 times, with the amount of amplified DNA doubling after each cycle. One of  
30 the PCR (oligonucleotide pair NO2.2 and CO3.3) resulted in the amplification of a single DNA fragment approximately 650 bp in length. Approximately 3 µg of the 650 bp fragment was generated by the PCR, representing about a  $5 \times 10^4$  fold amplification. The PCR was repeated using all  
35 combinations of the NO4 and CO3 series of probes. Based on the previous result, the combination of NO4.2 and CO3.3

would be expected to produce the amplified 650 bp fragment, and in fact this was the case.

**Example 3.**

**Identification of Cosmid Clones Carrying the Savinase-like**

**5 Alkaline Protease Gene From B. lentus**

To screen for the Savinase-like gene, the 1400 Tc<sup>R</sup> Km<sup>S</sup> HB101 clones were divided into 13 pools of 100-150 clones each. Plasmid DNA from each pool, and from a culture containing the entire gene bank was purified and digested  
10 with EcoRI. Restriction fragments were separated by agarose gel electrophoresis and transferred to Gene Screen Plus hybridization membrane (DuPont, Wilmington, Delaware) by the methods provided by the manufacturer. In order to detect the Savinase-like gene within the cosmid bank,  
15 oligonucleotide probes were synthesized. The predicted sequence of the probes was based on previously acquired amino acid sequence data for the Savinase protease. These probes also served as primers for the synthesis of a 650 bp portion of the BLAP gene using the PCR (described above).  
20 The 650 bp PCR fragment was labelled with <sup>32</sup>P in a nick translation reaction (protocol obtained from NEN, Wilmington, Delaware) and used as a hybridization probe. This fragment was shown to contain an internal EcoRI site. Therefore, two bands of hybridization would be detected  
25 when a complete B. lentus Savinase - like protease gene was present within a cosmid pool digested by EcoRI. When cosmid DNA isolated from the total gene bank of DSM 5483 DNA was digested with EcoRI, two fragments of 6.35 and 2.64 kb were found to hybridize with the labeled 650 bp PCR  
30 probe. In six of the recombinant cosmid pools (#3, #4, #8, #11, #12 and #13), these same bands of hybridization were detected, indicating that the Savinase-like gene was present in at least one cosmid clone contained within each of these six pools. Pool #13, which  
35 possessed 150 E. coli transformants, was subdivided into 12 subpools of 12 or 13 transformants each. Plasmid DNA from each subpool was isolated, digested with EcoRI, and the DNA

fragments separated by gel electrophoresis. After transfer to a Gene Screen Plus membrane, the filter was hybridized to the labeled 650bp PCR fragment. Two subpools (#7 and #9) were found to contain the same bands of hybridization as detected initially in Pool #13. Plasmid DNA was isolated from each individual transformant in subpools #7 and #9 and the plasmid DNA's were digested with either EcoRI or HindIII in separate reactions. After agarose gel electrophoresis and transfer to Gene Screen Plus, the hybridization procedure was repeated. Cosmid DNA from subpool #9 clone #5 exhibited the same pattern of hybridization expected of a clone which contained the Savinase-like protease gene. When cosmid DNA from clone #5 and DSM 5483 chromosomal DNA were digested with HindIII and probed with the 650bp PCR fragment, the same band of hybridization was detected for both. A fragment size of 3.4 kb was calculated from the location of the hybridization band in relation to known molecular weight standards. The cosmid from subpool #9 clone #5 was designated pSH44.

A large scale preparation of the cosmid pSH44 was prepared by CsCl- ethidium bromide gradient centrifugation, and after dialysis, restricted with HindIII. The HindIII fragments were separated by agarose gel electrophoresis, and the 3.4 kb HindIII fragment which hybridized with the 650bp PCR fragment was purified, and subcloned into the commonly used E. coli vectors pUC19 and pBR322 for restriction analysis and further subcloning into Bacillus vectors.

30

#### Example 4

**Genetic Characterization of the B. lentus Alkaline Protease Gene A. Subcloning and Characterization of a Restriction Fragment Carrying the B. lentus Alkaline Protease Gene**

After purification, the 650bp fragment was labeled by nick translation and used to probe gene banks of DSM 5483 DNA constructed using the cosmid pCP13 and EMBL lambda vectors. In both cases, using very stringent hybridization

conditions positive clones were identified. From hybridization data, several cosmid clones carrying DNA with homology to the 650 bp PCR fragment and selective oligonucleotide probes were identified. Plasmid DNA of one of these cosmids, pSH44, was purified and characterized by restriction analysis. Results from Southern blots indicated that a DSM 5483 HindIII fragment of approximately 3 to 5 kb hybridized to radioactive probes prepared from the 650 bp PCR fragment. Cosmid pSH44 carried three HindIII restriction fragments of 3.4 kb, 4.0 kb, and 4.8 kb. Initially, all of these fragments were purified by agarose gel electrophoresis and cloned into both pUC19 and pBR322. The ligation reaction in all cases contained 100 ng of dephosphorylated vector DNA and approximately 1000 ng of the corresponding HindIII fragment in a volume of 20  $\mu$ l. These ligation mixtures were transformed into competent cells of either E. coli DH5 $\alpha$  (for pUC19), or E. coli HB101 (for pBR322). Ampicillin resistant transformants were selected and screened for the correct plasmid constructs. All of the HindIII fragments were cloned into both vectors. The correct constructs in pUC19 and pBR322 were designated pMG102.2, and pMG105.9, respectively. The results from hybridization and restriction analysis identified the 3.4 kb HindIII fragment (present in plasmids pMG102.2 and pMG105.9) as having homology to the 650 bp PCR fragment. Plasmid pMG102.2 was purified, and characterized by restriction analysis with 26 different enzymes. In order to obtain the complete nucleotide sequence of the 3.4 kb HindIII fragment, it was necessary to subclone portions of the fragment into the pTZ sequencing vectors. All subclones were made using DNA fragments of pMG102.2 purified from low melting temperature agarose after gel electrophoresis. The excised DNA fragments were ligated to vector DNA which had been restricted at the appropriate sites, and transformed into E. coli DH5 $\alpha$  cells. Ampicillin resistant transformants were selected and the correct clones identified by restriction analysis. Plasmid DNA's

from the correct constructs were purified from the DH5 $\alpha$  strain and transformed into E. coli NM522 to provided the single stranded templates necessary for DNA sequencing.

**B. Sequencing of the 3.4 kb HindIII Fragment**  
**5 Containing the Alkaline Protease Gene of B. lentus**

The 650 bp PCR fragment was also cloned into E. coli plasmid pUC19. The PCR fragment was blunt-end ligated to pUC19 previously digested with SmaI, and transformed into E. coli DH5 $\alpha$  purchased from Bethesda Research Labs (BRL),  
10 using the procedure provided by BRL. Transformants were selected on X-gal plates containing 50  $\mu$ g/ml ampicillin (Maniatis, et al., ibid). Cloning of the 650 bp fragment into the SmaI site would be expected to disrupt lacZ complementation, and therefore produce white colonies on  
15 the X-gal plates. Approximately 200 of 10000 colonies on the X-gal plates were white. Plasmid DNA was prepared from 12 of the 200 transformants and analyzed for the presence of the 650 bp fragment. Seven of the 12 had the PCR fragment cloned into pUC19 in the same orientation. This  
20 plasmid was designated pMG100.13. For DNA sequencing, the 650bp fragment was removed from pMG100.13 by double digestion with BamHI and KpnI. Recognition sites for these enzymes are contained in the polylinker sequence originally from pUC19 but are not found within the DNA sequence of the  
25 650bp PCR fragment. The BamHI/KpnI PCR fragment was ligated with the commercial (Pharmacia) sequencing vectors pTZ18R and pTZ19R, digested with the same enzymes, and transformed into E. coli. Transformants containing the correct constructs were identified by isolation of plasmid  
30 DNA and restriction analysis. The two new constructs represented both orientations of the PCR fragment in relation to a sequencing primer site located on the pTZ vectors. DNA sequencing was performed using a dideoxy procedure and a commercially available sequencing primer  
35 from the M-13 system. The initial sequencing data consisted of approximately 220 bases for each strand, beginning at either end and extending internal into the 650

bp PCR fragment. This initial sequence confirmed that: the 650bp PCR fragment had been cloned into the SmaI site of pUC19 as a blunt end fragment; the DNA sequence immediately following the SmaI site in the pTZ18 clone is  
5 consistent with the NO2.2 primer; the 17 bp of DNA sequence immediately following the SmaI site in the pTZ19 clone (opposite orientation to pTZ18) is in agreement with the CO3.3 primer; and the amino acid sequences deduced from the DNA sequences both within and extending out from  
10 the NO2.2 and CO3.3 primer regions were in agreement with the N-terminal and internal amino acid sequences determined from Savinase. Based on the above results, it can be concluded that the 650bp PCR fragment amplified from B. lentus DSM 5483 DNA represents a portion of a gene encoding  
15 an alkaline protease, beginning just after the N-terminus of the mature protease and continuing into the protein coding region. To complete the sequencing of the 650 bp PCR fragment, two approaches were pursued. First, the pTZ19R vector carrying the 650 bp PCR fragment was digested  
20 with EcoRI to remove a portion of the PCR fragment, religated, and transformed into E. coli. Preliminary restriction analysis of the PCR fragment had revealed an internal EcoRI site, with a second EcoRI site present in the polylinker region of the vector. The correct subclone  
25 isolated from this experiment contained the EcoRI site adjacent to the priming site for DNA sequencing, allowing the determination of additional internal DNA sequence of the PCR fragment. At the same time, the small EcoRI fragment removed from the pTZ19R PCR construct was cloned  
30 in the opposite orientation into pTZ18R allowing the DNA sequence in the opposite direction from the EcoRI site to be determined.

**C. Determining the Transcriptional Start Site for the  
B. lentus 5483 Alkaline Protease Gene**

35 A necessary step in the genetic manipulation of the cloned BLAP gene is the determination of the start site of transcription. A primer extension protocol was used to

identify this site (Arnosti and Chamberlain (1989), PNAS(USA), 86, 830). In this procedure, a primer (20 to 30 bases in length) was synthesized which had homology to a sequence on the template DNA positioned 200 to 300 bp 3' to the suspected start of transcription. The primer was hybridized with the messenger RNA transcript under stringent conditions, and then the primer was elongated using the enzyme reverse transcriptase. The size of the reverse transcript was measured in a denaturing polyacrylamide gel. The elongated primer was also sequenced so that the exact base of transcription initiation could be read off of the sequencing gel. Total RNA was isolated as described by Arnosti and Chamberlin (ibid) from DSM 5483 or from the B. subtilis strain DB104 containing plasmid pJW5. Plasmid pJW5 was formed by cloning the 3.4 kb BLAP HindIII fragment into the HindIII site of plasmid pC194. Lysozyme treatment followed by sonication was used to disrupt Bacillus cells. To purify the RNA from contaminating nucleases, at least 4 to 5 phenol extractions were performed. The size of the BLAP gene transcript was determined by Northern analysis, using the protocol described by Thomas ((1983), Methods in Enzymology 100, 255). In this experiment, total Bacillus RNA was denatured and then analyzed by electrophoresis through an agarose gel containing a denaturing concentration of formaldehyde (BRL Focus 9:3,14). The RNA was then transferred from the gel to a nitrocellulose filter which was then baked at 60°C for 2 hours. Plasmid pMG109.3 (pUC19 + the 3.4 kb HindIII fragment containing the BLAP gene) was nick-translated and used to probe the Northern blot. The transcript produced by the BLAP gene was approximately 1200 to 1400 bases in length. Two oligonucleotide primers were designed on the basis of the BLAP nucleotide sequence. These primers were homologous to BLAP sequences between the NcoI and the HpaI sites, and 5' to of the sequence encoding the mature protease peptide. Primer A spans the BLAP nucleotide sequence included in Figure 2 between bases 2056-2033, and



primer B is found between bases 2242-2222. The results of the primer extension experiments using either primer and messenger RNA prepared from either B. lentus DSM 5483, or B. subtilis DB104 carrying plasmid pJW5 was the same. In both cases, the start site of transcription was determined to be an adenine residue at position 1924 in the BLAP sequence. This adenine is located 42 bp upstream of the ATG translation initiation codon at the beginning of the BLAP signal peptide. On the basis of this result, the length of the BLAP message is 1210 bases.

#### Example 5.

Construction of Fusions Between the B. licheniformis ATCC 53926 Alkaline Protease Gene and the B. lentus Alkaline Protease Gene (BLAP)

##### A. Experimental Approach

For large-scale production of BLAP, an expression system was designed which was based on gene fusions comprising parts of the ATCC 53926 alkaline protease gene and parts of the BLAP gene. The native promoter of the BLAP gene proved to be insufficient for the effective production of BLAP in B. subtilis DB104 as well as DSM 5483. It was known from fermentation results that the promoter of the ATCC 53926 alkaline protease gene worked well for the large scale production of Carlsberg-type protease in ATCC 53926. Thus, a fusion of the promoter of the ATCC 53926 alkaline protease gene to the structural gene of BLAP potentially would provide an efficient expression system for protease similar to commercially available serine proteases in ATCC 53926. Genes coding for alkaline proteases such as ATCC 53926 and BLAP typically encode a pre-pro form of the enzyme. They all start with a regulatory region comprised of a promoter, ribosome binding site and initiation codon, which is responsible for the efficient transcription and translation of the corresponding gene. The regulatory region is followed by a DNA sequence coding for the signal peptide. The signal sequence which is required for the secretion of the

protease outside the cell is followed by a coding region for pro-peptide and the mature protease. After transcription of the DNA sequence downstream from the promoter, the messenger RNA (mRNA) is translated into the final gene product. The pre-pro-mature protease consists of the signal peptide, the pro peptide and the mature protease. The signal sequence directs the immature protease through the cytoplasmic membrane, where the signal peptide is cleaved off enzymatically by a signal peptidase. The pro sequence seems to be essential for a correct folding and processing of the pro-protease during or after translocation across the cytoplasmic membrane and the cell wall. Amino acid sequence comparison between the BLAP and ATCC 53926 alkaline protease gene revealed 49% homology in the pre region, 38% homology in the pro region and 75% homology in the mature region. Because the interaction between the pre and pro regions necessary to obtain correct secretion and maturation of a protease is not fully understood, two different fusions between portions of the ATCC 53926 alkaline protease gene and portions of the BLAP gene were constructed. The two basic fusions were identified by the restriction enzyme sites used to join different portions of the ATCC 53926 and BLAP genes together.

#### 1. ClaI Fusion

In order to construct the ATCC 53926 alkaline protease-BLAP ClaI gene fusion, an extra ClaI restriction site had to be introduced into the ATCC 53926 alkaline protease gene at position 376 next to the junction of the ATCC 53926 pre and pro sequence (Figures 27 and 27a). At basepair 2032 the BLAP gene already contained a naturally occurring ClaI site at the corresponding position. The ClaI site was introduced into the ATCC 53926 alkaline protease gene by using the PCR. Plasmid pGM15 described previously was isolated from NM522, a *dam*<sup>+</sup> *E. coli* strain and restricted with AvaI, ClaI and MluI yielding four DNA fragments of 4330 bp, 1296 bp, 382 bp and 147 bp. The ClaI

sites at positions 1206 and 2032 in the BLAP sequence (Figure 2) were not restricted due to dam-methylation at overlapping GATC sequences. Restriction with MluI was used in order to generate three DNA fragments out of the 1825 bp

5 AvaI/ClaI fragment and to prevent religation and recloning of these fragments to the 4330 bp AvaI/ClaI fragment. The DNA fragments from the pGM15 AvaI/ClaI/MluI digest were ligated to a DNA fragment containing the ATCC 53926 alkaline protease promoter and pre sequence, which was

10 synthesized using the PCR. This 292 bp fragment was flanked by AvaI and ClaI restriction sites at its 5' and 3' ends, respectively, and was synthesized using plasmid pC51 as template with primers #1 and #3 (Figure 8). The PCR fragment contained sequence information from the N-terminal

15 part of the ATCC 53926 alkaline protease gene including the ATCC 53926 alkaline protease promoter, and most of the ATCC 53926 pre sequence (from the AvaI site at basepair 84 and extending to the newly formed ClaI site at basepair 376). The above ligation was transformed into E. coli NM522, and

20 ampicillin-resistant transformants were selected. Plasmid DNA's were prepared and characterized by restriction analysis. A correct construct was identified and designated pCB10C (Figure 12). Plasmid pCB10C contained the ATCC 53926 alkaline protease promoter and most of the

25 ATCC 53926 pre sequence fused to the BLAP sequence downstream from the native ClaI site at position 2941. Plasmid pCB10C was used to generate the final ClaI-fusion gene by linearizing it with ClaI and inserting a 909 bp ClaI fragment from the BLAP gene which contained the

30 sequence between the ClaI sites at position 2032, and position 2941. The 909 bp ClaI fragment was generated by hydrolyzing plasmid pMG108.2 (Figure 26) with ClaI. Plasmid pMG108.2 was isolated from E. coli GM33 (dam<sup>-</sup>) in order to prevent dam methylation at adenine residues in the

35 recognition tetranucleotide GATC. The 909 bp DNA fragment was isolated from a preparative agarose gel and ligated into plasmid pCB10C previously having been restricted with

ClaI and dephosphorylated with bacterial alkaline phosphatase. The ligation mix was transformed into E. coli NM522. Ampicillin resistant transformants were screened and the correct construct (pCB11C) was identified. A  
5 restriction map of pCB11C is given in Figure 4 and a schematic for its construction is given in Figure 12. Analysis of the DNA sequence across the fusion of the ATCC 53926 and BLAP genes present in plasmid pCB11C confirmed the expected DNA and amino acid sequences.

10        2. NcoI Fusion

In order to construct the ATCC 53926-BLAP NcoI gene fusion, an extra NcoI restriction site had to be introduced into the ATCC 53926 alkaline protease gene at position 638 near to the junction of the ATCC 53926 pro and mature  
15 sequence (Figures 27 and 27a). The BLAP gene contained a naturally occurring NcoI site at position 2311. The NcoI site was introduced into the ATCC 53926 sequence by using the PCR. A 554 bp PCR DNA fragment was synthesized using plasmid pC51 as template, and oligonucleotides #1 and #2  
20 (Figure 8) as primers. The design of the primers introduced AvaI and NcoI sites at the 5' and 3' ends of the fragment, respectively. The 554 bp fragment contained the sequence information for the ATCC 53926 alkaline protease promoter, pre and pro sequences, and the first four amino  
25 acids of the mature ATCC 53926 protease gene. After restriction with AvaI/NcoI the PCR fragment was ligated to the 4958 bp AvaI/NcoI DNA fragment from plasmid pGM15. Plasmid pGM15 (Figure 3) was restricted with AvaI, NcoI and MluI yielding four DNA fragments of 4958 bp, 668 bp, 382  
30 bp, and 147 bp. MluI was used to generate three fragments out of the 1197 bp AvaI/NcoI fragment to prevent religation of this fragment to the 4958 bp AvaI/NcoI fragment. The ligation mix was transformed into E. coli NM522. Ampicillin resistant transformants were screened and the  
35 correct construct (pCB1N) (Figure 5) identified. Analysis of the DNA sequence across the NcoI site of plasmid pCB1N confirmed the expected DNA and amino acid sequence. The

NcoI fusion alkaline protease gene contained the entire promoter, pre- and pro sequence of the ATCC 53926 and the entire sequence of the mature BLAP with a single amino acid substitution at amino acid residue 3 of the mature BLAP sequence where a threonine replaced a serine.

Example 6.

Subcloning of the B. licheniformis ATCC 53926 - BLAP Fusion Gene into Bacillus Vectors and Transformation into the ATCC 53926 Production Strain

After the ATCC 53926 alkaline protease-BLAP fusion genes had been cloned into E. coli vectors as previously described, they were subcloned into Bacillus vectors and transformed into B. subtilis DB104. The ATCC 53926 alkaline protease-BLAP fusion genes were subcloned as AvaI/SstI DNA fragments into pBC16 and pUB110 derivatives by restricting the E. coli vectors containing the ATCC 53926 alkaline protease-BLAP fusion genes (pCB1N, pCB11C, pCB2N, pCB13C, pCB4N and pCB15C) with AvaI and SstI, and ligating them to Bacillus plasmids pC51 and pH70 which had been restricted with AvaI, SstI, PstI and SalI. Plasmids pC51 and pH70 were cut with AvaI/SstI to provide the large AvaI/SstI vector fragment containing the necessary origin of replication and an antibiotic resistance gene (tetracycline resistance in pC51, kanamycin resistance in pH70) for selection. Restriction with SalI and PstI was done to hydrolyze the small AvaI/SstI fragment containing the entire ATCC 53926 alkaline protease gene present on pC51, and pH70 into three smaller DNA fragments. This decreased the probability of religating the ATCC 53926 alkaline protease gene into the Bacillus vectors. The initial cloning experiments into Bacillus were done using B. subtilis DB104 as recipient. Since DB104 is a reduced protease strain, it provides a direct screen for the presence and expression of the BLAP gene. Using pH70 and pC51 as recipients of the AvaI/SstI fragment from each of the E. coli, ATCC 53926 alkaline protease-BLAP fusion plasmids described above resulted in a total of 12

potential Bacillus plasmid constructs. Figure 23 depicts the six ATCC 53926-BLAP plasmids derived from the three ClaI fusions present in E. coli plasmids pCB11C, pCB13C, and pCB15C. Detailed restriction maps of the plasmids that were transformable into the ATCC 53926 production strain, plasmids pCB55C, pCB56C, pCB58C, pCB75C, and pCB78C are shown in Figures 13, 14, 15, 16, and 17 respectively. All five of these constructs were transformed into B. subtilis DB104 and into the ATCC 53926 production strain. In a similar fashion, the NcoI ATCC 53926 alkaline protease-BLAP fusions were cloned from pCB1N, pCB2N, and pCB4N into pH70 and pC51. Five of the six plasmids from this cloning (pCB50N, pCB51N, pCB53N, pCB70N, and pCB73N) were transformed into B. subtilis DB104, and into the ATCC 53926 production strain. The restriction maps for these constructs are shown in Figures 18, 19, 20, 21, and 22 respectively.

#### Example 7

**Protocol for Production of B. lentus Alkaline Protease (BLAP) by ATCC 53926 Alkaline Protease-BLAP Strains in Shake Flask Cultures.**

To compare derivatives of B. licheniformis ATCC 53926 carrying the different ATCC 53926 alkaline protease-BLAP plasmids, a three-stage shake flask procedure was used. In these experiments, the antibiotic tetracycline or kanamycin was always present in the media (described below) to select for the desired plasmid. The first preculture flask was inoculated from a frozen stock vial that had been thawed at room temperature. After 7 hours of incubation at 39°C and 200 rpm, 1.0% (v/v) of this culture was transferred to a fresh (second) preculture flask and the incubation resumed. Between 13 and 15 hours, a 1% transfer was made to duplicate main culture flasks. Culture samples were obtained periodically from the main culture flasks between 27 and 50 hours for determination of protease activity. The shake flask experiments used a complex medium that contained assimilable carbon and nitrogen sources for the

microorganism as well as other organic materials and inorganic salts. The major components included hydrolyzed corn starch, sodium caseinate, soy meal, a brewing residue, and corn steep liquor. Small amounts of antifoam, 5 inorganic salts and vitamins B1 and B6 were also present. After distribution to baffled erlenmeyer flasks the medium was sterilized by heating to 121°C at 15 psi.

#### Example 8

#### Protocol for Small-Scale Fermentation of ATCC 53926 10 Derivatives that Produce BLAP.

The fermentations were carried out in two steps where the organism was first grown up in a shake flask containing a complex medium with antibiotic to select for the endogenous plasmid DNA. After achieving a high cell 15 density, part of the flask culture was transferred to the main fermenter vessel for continued growth and the production of protease. The vessels were Biostat E or ES manufactured by B. Braun, Inc.. All fermentations contained less than 10 liters of culture, and were carried 20 out with BL1 (Biosafety Level One) containment precautions and with good laboratory practices as specified in the National Institutes of Health Guidelines for Research Involving Recombinant DNA Molecules (NIH Guidelines, May 7, 1986).

25 The shake flask and fermenter media were very similar complex media with an assimilable carbon source, assimilable nitrogen source, organic nutrients and inorganic salts which allowed vigorous growth of the microorganism as well as production of protease. Major 30 medium components included deionized water, corn starch, soy meal, sodium caseinate, a brewing residue, and corn steep liquor. Minor components were  $(\text{NH}_4)_2\text{HPO}_4$  and  $\text{Na}_2\text{HPO}_4$ . A commercial antifoam was also present in the fermenter media. The medium was sterilized by heating to 121°C at 15 35 psi. The described media in a two-liter baffled flask was inoculated from 1.0 ml of frozen culture-glycerol stock (see stock preparation) that had been thawed at room

temperature. Either tetracycline or kanamycin antibiotic was added to the shake flask culture to select for the cells that carried the desired plasmid. The shake flasks were incubated at 39°C with slow shaking to grow the culture. The fermenter media was eventually inoculated with a 1.5% (v/v) transfer of the shake flask culture. Antibiotic was also present in the fermenter media. The agitation rate and airflow were held at 1,300 rpm and 1 vvm (liters/min.), respectively. In a few experiments, the agitation and airflow rate had to be increased for a brief period during log phase growth to maintain a partial oxygen pressure ( $pO_2$ ) of greater than 20% in the culture.

During cultivation of the production strain, the culture pH and temperature were changed at specified times. In particular, it was found that control of the culture pH had a significant effect on production of protease. The culture pH was controlled at 6.8 from 0-13.5 hours, then controlled at pH 7.5 from 18 to 35 hours. Sterile solutions of KOH and  $H_2SO_4$  were added to the media as needed to control pH. From the start of the fermentation the culture temperature was 39°C. The temperature was lowered to 34°C after 14.5 hours of incubation.

#### Example 9

##### Assays for the B. lentus Alkaline Protease

For assays of samples from shake flasks or fermentations, the protease was separated from the cells and media debris by centrifugation. The protease in these crude preparations was quantified as described below.

The HPE (Henkel Protease Einheit) method was based on a discontinuous assay of the release of acid-soluble fragments from protease digests of casein. About 50  $\mu$ l of protease solution was combined with 750  $\mu$ l of a casein solution prepared by heating Hammarsten casein in tripolyphosphate-tris buffer solution at pH 8.5. After 15 minutes of incubation in a 2.2 ml eppendorf tube at 50°C, 600  $\mu$ l of 0.44M trichloroacetic acid solution was added. The mixture was chilled on ice for 15 minutes before



removing the precipitate by centrifugation. Acid-soluble peptides were measured in a spectrophotometer at 290 nm. Two HPE units were defined as being equal to an absorbance change of 0.500 OD at 290 nm.

5           The EPE (Esterolytic Protease Einheit) method was useful primarily for detection of activity from the ATCC 53926 production strain. The kinetic assay was started by mixing 100  $\mu$ l of a sample or known standard of a commercial protease (Maxatase, Novo A/S) into an acetone buffer  
10 containing CBZ-valine-p-nitrophenyl ester. The rate at which p-nitrophenol was released by esterolytic cleavage of the CBZ ester was monitored spectrophotometrically at  $A_{340}$  and determined by linear regression analysis. Units of EPE activity for a sample were determined by comparison of the  
15 reaction rate of the sample to that of the Maxatase standard.

#### Example 10.

##### Distinguishing between BLAP and B. licheniformis ATCC 53926 Protease by Gel Electrophoresis

20           The B. lentus alkaline protease produced by the ATCC 53926 alkaline protease-BLAP fusion strains could be distinguished from ATCC 53926's native extracellular proteolytic enzymes by electrophoresis in homogeneous, non-denaturing ("native") 12.5% polyacrylamide (PA) gels with  
25 pH 6.0 buffer. Crude samples obtained by centrifugation of the fermenter broth were resolved by this method on the PhastSystem apparatus manufactured by Pharmacia, Inc.. The protein species were visualized by staining the PA gel in Coomassie Blue.

30

#### Example 11

##### Preparation of B. licheniformis ATCC 53926

Cell material from agar slants containing strain B. licheniformis DSM641 was transferred by a loop and suspended in 5 ml of medium A (Medium A: Nährbouillon from  
35 Merck prepared according to the manufacturer's directions). After intensive agitation in a Vibrox the suspension was agitated at 30°C in a circular shaker (150 RPM, amplitude

of 2.5 cm; 24 hr). After 24 hours, 1 ml of the suspension (Optical Density = 1 at 640 nm) was incubated in 100 ml of medium A in a 500 ml Erlenmeyer shake flask for an additional 24 hours. Then it was centrifuged off at 6500 rpm. The centrifugate was rinsed twice with an isotonic NaCl solution and placed in enough salt water to give a cell count of approx.  $10^9$ /ml.

About 8 ml of the suspension were transferred to a 9 cm diameter Petri dish for U.V. radiation (5-60 minutes) at a distance of 27 cm from the U.V. lamp (Schutt/Göttingen) while agitating at 100 rpm. The agitation was continued during radiation to prevent settling of the cells. After the desired radiation time, preferably 6 minutes, a dilution series was made and these dilutions were plated on medium-A agar plates with and without antibiotic. From the determination of CFU (Colony Forming Units) before and after mutagenic treatments, with and without antibiotics, the mortality rate can be determined as well as changes in reversion rate which serves as a measurement for mutation formation. In the case of DSM 641, a mortality rate of >95% was considered to be especially desirable.

Using a replica plating technique, the colonies were transferred to Ca-caseinate-agar. The plates were incubated at 39°C for 72 hours. Then at 24, 48, and 72 hours the formation of clearing zones around the colonies was measured. After measuring the colonies and the clearing zones, the strains were selected which showed a large lysis area as compared with the diameter of the colony. Then individual cultures of these colonies were produced using the standard microbiological procedures (Koch's Agar Plate Procedure, Lindner's Drop Procedure, Multiple Streak Procedure). Shake flask experiments were undertaken with these isolates to determine protease productivity.

## Example 12

## Comparison of Protease Production by DSM641 and ATCC 53926

The strains were incubated in shake flasks with 50 ml of medium B (see footnote 1 to Table III) in a 500 ml Erlenmeyer for 96 hours at 39°C and 150 RPM (amplitude of 5 cm). The protease activity was determined every 24 hours. The relative protease production is summarized in Table III.

Table III

10	Strain	Protease activity		Ratio		Average		Ratio		Avg
		EPE/ml		Rel. to		Protease		Rel. to		
		72 <sup>3</sup>	96	72	96	72	96	72	96	
15	53926 <sup>2</sup>	4950	6702	1.36	1.32					
	53926	5298	6340	1.45	1.35	5120	6270	1.35	1.33	
	DSM641	3950	4920	1.08	1.05					
	DSM641	3650	4690	1.00	1.00	3800	4805	1.00	1.00	

1. Medium B:
 

KH <sub>2</sub> PO <sub>4</sub>	2.5 g/l
MgSO <sub>4</sub> ·7H <sub>2</sub> O	1.0 g/l
MnSO <sub>4</sub> ·2H <sub>2</sub> O	0.5 g/l
CaCl <sub>2</sub> ·2H <sub>2</sub> O	0.2 g/l
Soy bean meal	3.0 g/l
Casein nach Hammersten	15.0 g/l
Potato starch	120.0 g/l
Amylase	0.6 g/l
2. Duplicate fermentations were performed.
3. The entries 72 and 96 refer to incubation time in hours.

Table III shows that the strain ATCC 53926 produces about 30% more protease compared to strain DSM641.

### Miscellaneous Experimental Operations

#### 1. Preparation of Plasmid DNA

Plasmid preparation (mini-prep. and CsCl/ethidium bromide density gradient purified), agarose and acrylamide gel electrophoresis, buffers (TE- Tris EDTA, TAE- Tris acetate EDTA, TBE- Tris borate EDTA) and media preparation were essentially performed as described by Maniatis, et al., (ibid). Phenol p.a. was purchased from IBI and equilibrated with TE-buffer prior to use. Agarose (SeaKem GTG) and low gelling temperature agarose (SeaPlaque) were purchased from FMC Corp., Rockland, ME.

#### 2. Purification of Restriction Fragments

Restriction fragments were isolated by running the sample on a preparative agarose gel. DNA was visualized by staining with ethidium bromide and the band with the DNA fragment to be recovered cut out with a razor blade. The DNA was electroeluted using an IBI electroeluter, phenolized once and precipitated with ethanol.

#### 3. Polymerase Chain Reaction

The PCR was performed exactly as described by the manufacturer (Cetus Corporation, Emeryville, California). The solution containing the DNA generated by PCR was extracted with phenol, precipitated with ethanol and resuspended in an appropriate buffer for hydrolysis with restriction enzymes. After hydrolysis with the appropriate restriction enzyme, the solution was extracted with phenol and the DNA was precipitated with ethanol. DNA prepared in this manner was suitable for ligation. In this report the expected length of all PCR generated DNA fragments corresponded to the size of the DNA fragments observed after restriction with the appropriate restriction enzymes.

#### 4. Restriction Analysis

Restriction enzymes, T4 DNA Ligase and bacterial alkaline phosphatase were used as described by the manufacturers (NEB, BRL, IBI, Boehringer).

### 5. Protoplast Transformation

The method of protoplast transformation originally optimized for B. subtilis was adapted to other Bacillus species (Chang and Cohen (1979) Mol. Gen. Genetics 168, 111). A colony of B. subtilis, B. lentus, or B. licheniformis was inoculated into 30ml of 416 media (416 media: 2% Bactotryptone, 1% Yeast Extract, 1% NaCl, sterilized by autoclaving) and grown overnight at 37°C in a New Brunswick G24 tabletop shaker at 225 rpm. The next morning, an appropriate volume of each culture was used to inoculate 40 ml of 416 broth (prewarmed) in a sterile 250ml Erlenmeyer flask. Incubation at 37°C was continued with shaking, until an OD<sub>650</sub> between 0.4 and 0.5 was reached. Then the cells were pelleted in a sterile 35ml polypropylene Oakridge type tube, 3000rpm, 10min at 4°C. The supernatant was removed by pouring, and the cell pellet gently resuspended in 5ml of SMMP by inversion. SMMP was always prepared fresh by mixing equal volumes of 2X SMM and 4X PAB (2X SMM contained: 1.0M Sucrose, 0.04M Maleic Acid, 0.04M MgCl<sub>2</sub>·6H<sub>2</sub>O, pH6.5 adjusted with 10%NaOH, sterilized by autoclaving; 4X PAB contained: 70 g of Difco Antibiotic Medium #3 dissolved in one liter of deionized water and sterilized by autoclaving). The cell suspension was transferred by pouring into a sterile 16 x 125mm round bottom, screw-cap culture tube (Corning #25760). A lysozyme solution, 10mg/ml in SMMP was prepared fresh, and 200 µl added to the cell suspension. The tubes were placed on the platform of a G24 New Brunswick shaker, and incubated at 37°C with very gentle shaking (25-30 rpm). Protoplast formation generally took 40-60 minutes for B. licheniformis, and 20-40 minutes for B. lentus and B. subtilis. To determine the degree of protoplasting, a sample of the cell suspension was examined using the phase contrast microscope. Following the lysozyme treatment, 10ml of SMMP was added to the protoplast suspension and the protoplasts were pelleted by centrifugation at 2000 rpm for 10 minutes at room temperature. The supernatant was

carefully removed before resuspending the protoplast pellet in SMMP (3ml of SMMP for each 40 ml of initial cell culture). In each transformation reaction, 0.5ml of protoplasted cells was transferred into a sterile, plastic 5 16 x125mm round bottom culture tube (Corning #256760), followed by the addition of one to four micrograms of DNA in a volume of 60 $\mu$ l or less, and 1.5 ml of 40% polyethylene glycol (PEG, Sigma Chemical Company product #P-3640). After gentle mixing of the tubes by inversion they were 10 allowed to sit at room temperature for 3-4 minutes, and then the PEG was diluted by the addition of 5ml of SMMP + 2% BSA (Bovine Serum Albumin). After mixing by inversion, the protoplasts were pelleted by centrifugation at 2000 rpm for 10 minutes at room temperature, and the supernatant was 15 carefully poured off before resuspending the pellet in 1.0ml of SMMP+BSA. The cells were then incubated at 37°C for 90 minutes with shaking (New Brunswick G24 table top shaker, 150 rpm) to allow for expression of plasmid encoded genes, and to partially regenerate cell walls. Selective 20 regeneration medias were prepared with the appropriate antibiotic. The CR5 regeneration medium was used for all Bacillus strains when selecting for transformants containing plasmids encoding Km<sup>R</sup> (Puyet et al. (1987) FEMS Microbiology Letters 40, 1). The DM3 regeneration medium 25 was originally used to select all Bacillus strains transformed by plasmids encoding Tc<sup>R</sup> (Chang and Cohen (ibid)). However, B. lentus protoplasts were discovered to regenerate inefficiently on this succinate based media. Therefore, #451 regeneration medium which used mannitol as 30 an osmotic stabilizer was tested (Bourne and Dancer (1986) J. of Gen. Micro. 132, 251). All of the Bacillus strains were found to regenerate well on the mannitol media, although the level of tetracycline needed for each strain was different. The selective levels of tetracycline for B. subtilis, B. lentus, and ATCC 53926 were 65  $\mu$ g/ml, 15 35  $\mu$ g/ml, and 25  $\mu$ g/ml, respectively. Each transformation reaction was plated onto 5 regeneration plates and

incubated 1-2 days at 37°C. Potential transformants were picked onto Luria Agar containing the appropriate antibiotic, and when necessary 1% skim milk to detect protease activity. To check strain identity, Spizizens's  
5 minimal salts medium with and without the correct nutritional additives were used.

#### 6. Bacterial Strains

##### Escherichia coli Strains:

1. E. coli NM522 is hsdA5, Δ(lac-pro), [F', pro<sup>+</sup>,  
10 lacI<sup>q</sup>ZAM15]. NM522 is restriction minus and modification minus (Pharmacia, Piscataway, N.J.)
2. E. coli GM33 (dam3) is dam-methylase minus (Marinus, M. G. and Morris, N. R., (1974) J. Mol. Biol. 85, 309).
3. E. coli DH5α is F-φ80dlacZAM15, Δ(lacZYA-argF)U169,  
15 recA1, endA1, hsdR17(r<sub>k</sub><sup>-</sup>, m<sub>k</sub><sup>+</sup>), supE44, lambda<sup>-</sup>, thi-1, gyrA, relA1 (competent frozen cells, BRL, Gaithersburg, MD).
4. E. coli HB101 is F<sup>-</sup>, hsdS20 (r<sub>B</sub><sup>-</sup>, m<sub>B</sub><sup>-</sup>), supE44, ara14, glaK2, lacYi, proA2, rpsL20(str<sup>R</sup>), xyl15, leu, mt11, lambda<sup>-</sup>, recA13.

##### 20 7. Bacillus Strains:

1. B. subtilis DB104 has the genotype: his, nprR2, nprE18, and aprA3. This strain was obtained from Dr. Roy Doi at the University of California at Davis through a licensing agreement, and has only been used for research purposes.  
25 This strain is deficient in the production of neutral and alkaline protease.
2. B. licheniformis ATCC 53926 is an industrial production strain developed by Henkel KGaA, Düsseldorf, West Germany.
- 30 3. B. lentus DSM 5483 is a natural isolate of a soil sample from West Germany.

#### 8. Cloning Vectors

1. pUC19: pUC19 is a small E. coli plasmid (2686 bp) containing parts of pBR322 and M13mp19. It carries a 54 bp  
35 multiple cloning site with recognition sites for 13 different restriction endonucleases. It has been constructed by C. Yanisch-Perron et al. ((1985), Gene 33,

103), and can be purchased from New England Biolabs, Inc., Beverly, MA.

2. pTZ18R: pTZ18R (Mead, D.A. *et al.* (1986) Protein Engineering, 1, 67) is a multifunctional phage plasmid (phasmid) vector. It contains origins of replication of the E. coli vector pBR322 and of the E. coli phage f1, allowing single-stranded DNA to be produced. It contains the multiple cloning site of plasmid pUC18. It can be purchased from Pharmacia LKB Biotechnology, Inc., Piscataway, NJ.
3. pTZ19R: pTZ19R is homologous to pTZ18R but contains the multiple cloning site of plasmid pUC19.
4. pCP13: pCP13 is a broad host range cosmid vector derived from plasmid pLAFR1, which has been previously characterized (Friedman, *et al.*, (1982) Gene 18, 289). Plasmid pCP13 is 23 kb in size, and encodes resistance to tetracycline and kanamycin. It contains multiple unique restriction sites suitable for cloning, can be mobilized, but is not self-transmissible.
5. pC50: pC50 is a derivative of the Bacillus plasmid pBC16. It contains a 2 kb DNA fragment encoding the DSM641 protease cloned into the BamHI site of pBC16.
6. pC51: pC51 is a second generation derivative of plasmid pC50. The 1540 bp AvaI/SstI fragment from pC50 which encodes an ATCC 53926 alkaline protease, was first cloned into the E. coli plasmid pUC19 between the XmaI and SstI sites to form plasmid pH9. The ATCC 53926 protease gene was then removed from pH9 on an EcoRI/BamHI fragment and cloned into plasmid pBC16 between the EcoRI and BamHI sites to form plasmid pC51.
7. pUB110: pUB110 is a 4548 bp plasmid encoding resistance to kanamycin, which was originally isolated from Staphylococcus aureus by R.W. Lacey ((1974) J. Med. Microbiol., 7, 285).
8. pKO110: pKO110 is a derivative of pUB110 in which the 788 bp EcoRI/BamHI fragment has been removed, and replaced by a fragment from pUC19 containing the polylinker region.



9. pBC16: pBC16 has an approximate size of 4250 bp, and encodes resistance to tetracycline. This plasmid was originally isolated from Bacillus cereus and transformed into B. subtilis (Bernhard, K., et. al., (1978) J. Bacteriol. 133:897).
- 5 10. pC194: pC194 is a 2910 basepair plasmid encoding resistance to chloramphenicol which was isolated from S. aureus and transformed into B. subtilis. (Iordanescu, S., et. al., (1978) Plasmid 1:468) and (Erlich, S.D. (1977), PNAS USA 74:1680).
- 10 11. pH70: pH70 is a derivative of plasmid pUB110. The ATCC 53926 alkaline protease gene was removed from plasmid pH9 (see plasmid pC51) on an EcoRI/BamHI DNA fragment and cloned into plasmid pUB110 between the EcoRI and BamHI sites to form plasmid pH70.
- 15

What is claimed is:

1. A composition comprising a mature alkaline proteolytic enzyme from Bacillus lentus DSM 5483 having the amino acid sequence from about amino acid residue 112 to about amino acid residue 380 substantially as shown in Figure 1 and at least one other proteolytic enzyme from Bacillus licheniformis ATCC 53926.
2. A hybrid plasmid capable of replicating in Bacillus containing a DNA sequence comprising in the direction of transcription a regulatory region, said regulatory region comprising in the order of transcription a Bacillus licheniformis ATCC 53926 alkaline protease promoter, ribosome binding site, and initiation codon operably linked to a pre-pro region, said pre-pro region operably linked to a mature region encoding for said proteolytic enzyme from Bacillus lentus DSM 5483 of claim 1.
3. The hybrid plasmid of claim 2 further comprising a Bacillus licheniformis ATCC 53926 alkaline protease transcription terminator operably linked to said mature region encoding for said proteolytic enzyme from Bacillus lentus DSM 5483 of claim 1.
4. The hybrid plasmid of claim 2 further comprising a carboxy terminal portion of the Bacillus licheniformis ATCC 53926 alkaline protease gene and its transcriptional terminator operably linked to said mature region encoding for said proteolytic enzyme from Bacillus lentus DSM 5483 of claim 1.
5. An isolated DNA sequence coding for the pre-pro-mature form of said proteolytic enzyme from Bacillus lentus DSM 5483 of claim 1.

6. The hybrid plasmid pCB55C having the cleavage map substantially as shown in Figure 13.
7. The hybrid plasmid pCB56C having the cleavage map substantially as shown in Figure 14.
8. The hybrid plasmid pCB58C having the cleavage map substantially as shown in Figure 15.
9. The hybrid plasmid pCB75C having the cleavage map substantially as shown in Figure 16.
10. The hybrid plasmid pCB78C having the cleavage map substantially as shown in Figure 17.
11. The hybrid plasmid pCB50N having the cleavage map substantially as shown in Figure 18.
12. The hybrid plasmid pCB51N having the cleavage map substantially as shown in Figure 19.
13. The hybrid plasmid pCB53N having the cleavage map substantially as shown in Figure 20.
14. The hybrid plasmid pCB70N having the cleavage map substantially as shown in Figure 21.
15. The hybrid plasmid pCB73N having the cleavage map substantially as shown in Figure 22.
16. A transformed Bacillus host cell comprising a hybrid plasmid of claim 6.
17. A transformed Bacillus host cell comprising a hybrid plasmid of claim 7.

18. A transformed Bacillus host cell comprising a hybrid plasmid of claim 8.
19. A transformed Bacillus host cell comprising a hybrid plasmid of claim 9.
20. A transformed Bacillus host cell comprising a hybrid plasmid of claim 10.
21. A transformed Bacillus host cell comprising a hybrid plasmid of claim 11.
22. A transformed Bacillus host cell comprising a hybrid plasmid of claim 12.
23. A transformed Bacillus host cell comprising a hybrid plasmid of claim 13.
24. A transformed Bacillus host cell comprising a hybrid plasmid of claim 14.
25. A transformed Bacillus host cell comprising a hybrid plasmid of claim 15.
26. The transformed Bacillus host cell of claim 16 wherein said host cell is Bacillus licheniformis ATCC 53926.
27. The transformed Bacillus host cell of claim 17 wherein said host cell is Bacillus licheniformis ATCC 53926.
28. The transformed Bacillus host cell of claim 18 wherein said host cell is Bacillus licheniformis ATCC 53926.
29. The transformed Bacillus host cell of claim 19 wherein said host cell is Bacillus licheniformis ATCC 53926.

30. The transformed Bacillus host cell of claim 20 wherein said host cell is Bacillus licheniformis ATCC 53926.
31. The transformed Bacillus host cell of claim 21 wherein said host cell is Bacillus licheniformis ATCC 53926.
32. The transformed Bacillus host cell of claim 22 wherein said host cell is Bacillus licheniformis ATCC 53926.
33. The transformed Bacillus host cell of claim 23 wherein said host cell is Bacillus licheniformis ATCC 53926.
34. The transformed Bacillus host cell of claim 24 wherein said host cell is Bacillus licheniformis ATCC 53926.
35. The transformed Bacillus host cell of claim 25 wherein said host cell is Bacillus licheniformis ATCC 53926.
36. A method for producing the enzyme of claim 1 comprising cultivating a Bacillus licheniformis ATCC 53926 strain which has been transformed by a plasmid having a gene encoding for said enzyme in a nutrient medium and recovering said enzyme.
37. The method of claim 36 wherein said plasmid is pCB55C.
38. The method of claim 36 wherein said plasmid is pCB56C.
39. The method of claim 36 wherein said plasmid is pCB58C.
40. The method of claim 36 wherein said plasmid is pCB75C.
41. The method of claim 36 wherein said plasmid is pCB78C.
42. The method of claim 36 wherein said plasmid is pCB50N.
43. The method of claim 36 wherein said plasmid is pCB51N.

44. The method of claim 36 wherein said plasmid is pCB53N.
45. The method of claim 36 wherein said plasmid is pCB70N.
46. The method of claim 36 wherein said plasmid is pCB73N.
47. A composition comprising a mature alkaline proteolytic enzyme having the amino acid sequence substantially as shown in Figure 29 and at least one other proteolytic enzyme from Bacillus licheniformis ATCC 53926.
48. A biologically pure culture of Bacillus licheniformis having all the characteristics of the strain deposited under accession number ATCC 53926.
49. A biologically pure culture of Bacillus lentus strain having all the characteristics of the strain deposited under accession number DSM 5483.
50. A composition comprising an alkaline proteolytic enzyme of claim 1 in combination with a surfactant.
51. A composition of claim 50 wherein said surfactant is selected from the group consisting of a linear alkyl benzene sulfonate, an ethoxylated linear alcohol, an alkyl ethoxylated sulfate, a sulfated linear alcohol, or combinations thereof.

	5	10	15	20	25	30																								
1	M	K	K	P	L	G	K	I	V	A	S	T	A	L	L	I	S	V	A	F	S	S	S	I	A	S	A	A	E	E
31	A	K	E	K	Y	L	I	G	F	N	E	Q	E	A	V	S	E	F	V	E	Q	V	E	A	N	D	E	V	A	I
61	L	S	E	E	E	E	V	E	I	E	L	L	H	E	F	E	T	I	P	V	L	S	V	E	L	S	P	E	D	V
91	D	A	L	E	L	D	P	A	I	S	Y	I	E	E	D	A	E	V	T	T	M	A	Q	S	V	P	W	G	I	S
121	R	V	Q	A	P	A	A	H	N	R	G	L	T	G	S	G	V	K	V	A	V	L	D	T	G	I	S	T	H	P
151	D	L	N	I	R	G	G	A	S	F	V	P	G	E	P	S	T	Q	D	G	N	G	H	G	T	H	V	A	G	T
181	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	E	L	Y	A	V	K	V	L	G	A	D	G	R
211	G	A	I	S	S	I	A	Q	G	L	E	W	A	G	N	N	G	M	H	V	A	N	L	S	L	G	S	P	S	P
241	S	A	T	L	E	Q	A	V	N	S	A	T	S	R	G	V	L	V	V	A	A	S	G	N	S	G	A	S	S	I
271	S	Y	P	A	R	Y	A	N	A	M	A	V	G	A	T	D	Q	N	N	N	R	A	S	F	S	Q	Y	G	A	G
301	L	D	I	V	A	P	G	V	N	V	Q	S	T	Y	P	G	S	T	Y	A	S	L	N	G	T	S	M	A	T	P
331	H	V	A	G	A	A	L	V	K	Q	K	N	P	S	W	S	N	V	Q	I	R	N	H	L	K	N	T	A	T	
361	S	L	G	S	T	N	L	Y	G	S	G	L	V	N	A	E	A	A	T	R										

## Amino acid composition

-----

54 A	0 C	8 H	5 M	21 T
9 R	12 Q	19 I	6 F	3 W
24 N	28 E	30 L	17 P	9 Y
11 D	36 G	10 K	43 S	36 V

Number of residues: 380

Molecular weight (MW): 38951

**FIG. 1****SUBSTITUTE SHEET**

**FIG. 2-1**

10 20 30 40 50

1 AAGCTTCTAGAGATCTTCCATACCTACCAAGTTCTGCGCCTGCAGCAATGG  
51 CAACAACGTCGCGGATCATTACAAAGACGATTGCCAAGGCCGCTAACTGA  
101 ATACTGCTAGCTGTGTACCACTGTAAAGCGAAAGGTGCGTCTAAGTAATA  
151 AGGCCGAACCTTAAATTAATTCCTCGTTTGCCACACAGGCTTTTATAG  
201 GCGTGTGTATTTTGTATGCCTAGTAGAATCAAATAGACATAAGAACTG  
251 GAGGACTTATTATATGCATGAACAATTGAGACAAACGACACAACCAACGA  
301 TTATTGTTGGTGTGGAGTTCGAAAAAGACACGGATTTCCACTATTCGATG  
351 GAAGAATTGAACAATTTGGCCGATCGCTTGATTGTCAGGTTGTGCAACAGC  
401 TCACACAAAAGCTGCCTGTACCAAATCAAGCGACATATATTGGTGTGGC  
451 AAAGCGTCTGAGCTTTCCTGATGTGTGAATCGCTGGGGGCTACGCTCGT  
501 TGTGTTTAAATGACGAGCTCTCACCTTCACAAATCCGCAATCTTGAAAAAT  
551 TGCTGGAAGTAACCGTTTATGATCGGACCATGTTGATTTTAGATATTTT  
601 GGCGAAAGGGCGAAGACGAAAGAAGCTCAACTCCAAGTAGAAATGGCGCG  
651 CTGCCGCTATTTGTTGCGACGGCTCGTCGGCATGCGCGCTTCATTAAAGCA  
701 GGCAAGGCGGCCGGCACTGGCTTGCCAAACGCGGCGCGGGGAAACAAA  
751 GTTGGAGCTCGATCGTCGTAAAATTGAAACACGTATTTCATGCGTTGAAAA  
801 AAGAGCTGGAAGACGTTGTCAAACGACGCGAGTTACAGCGCAAACGCCGG  
851 AAAAAACAAGCGATGCCGGTCGTTGCCTTAGTTGGCTATACAAATGCAGG  
901 GAAATCTTCGCTGCTGAATGCGCTGTTGGACGATGCCGAGGAAAAAGCGTG  
951 TGTGGAAAAAGATATGTTGTTTGCCACTCTTGATACGTCTGTGCGTAAA  
1001 GTGGAACCTTGACAAAAACCATTTCTGTTTTATTAGCGGATACGGTTGGCTT  
1051 TGTCTCTAAATTGCCAACACACCTCGTTAAAGCATTCCGCTCAACATTGG  
1101 AAGAAAGCTCGGGAAGCCGATTTGCTACTGCATGTCGTCGATTATCAAAT  
1151 GAACGCCATCGCGAAATGGCAAAGACGACAAATGAAACACTCCAGGCAAT  
1201 GGAAATCGATCGCCGATGATTTATGTTTACAACAAAATGGATCGAGTGA  
1251 AAGACGCGTTTCTCAAGCGCATGGGCACGAGCTGTTTATATCAGCTAAG  
1301 GCTAAACAAGGGCTTGATTTATTAGCACAGAAAATAGCAAGCTATGTTTT  
1351 TCAAGATTTTGAAAAACATCTGTTTCATCATTCTTATCGTGACGGGGAGG  
1401 CGGCTGCTTATTTAAACAACCATGCCCATGTTACACACAGCGTGCTGAG  
1451 GAGGACGGCTGGCATATCGTTGCCGATTTGTCATGAACGAGACTTGAAACG  
1501 GGTTGAAAGCTACTGTGTTTCAAAGAACGATAATGAAAAAAGCCATTG  
1551 AATGCTTCTTGTTCAAATGGCTTTTTGGCGACTATGGTAGACAGATGAAC  
1601 ACTTGTTTCGCTGTTTTACGACAAAGATCATCCTGCCTGTTACGCGTTTT  
1651 TTAAATCCGTTTTCGCACGTTCAATTGTCGCCGAGTCGTACCAAGTCGCTG  
1701 TAAGTGAGAATATGTTTAGAAAGCCGCGTATTTAAGCGCAGTCTTTTTCG  
1751 TTCTGTACTGGCTGGTTTTGTGGACAGTTTCCATACCCATCGGCCTCCTTT  
1801 TATTTGTAGCTTTCCCACTTGAAACCGTTTTAATCAAAAAAGAAACGGG  
1851 AAGATTCTGTAACTTAACGTTAATATTTGTTTCCCAATAGGCAATCTT  
1901 TCTAACTTTGATACGTTTAACTACCAGCTTGGACAAGTTGGTATAAAAA  
1951 TGAGGAGGGAACCGAATGAAGAAACCGTTGGGAAAATTGTCGCAAGCAC  
2001 CGCACTACTCATTTCTGTTGCTTTTAGTTCATGCATCGCATCGGCTGCTG  
2051 AGGAAGCAAAAGAAAAATATTTAATTGGCTTTAATGAGCAGGAAGCTGTC  
2101 AGTGAGTTTGTAGAACAAGTAGAGGCAAATGACGAGGTCGCCATTCTCTC  
2151 TGAGGAAGAGGAAGTCGAAATTGAATTGCTTCATGAGTTTGAACGATTCT  
2201 CTGTTTTATCCGTTGAGTTAAGCCCAGAAGATGTGGACGCGCTTGAACCTT  
2251 GATCCAGCGATTTCTTATATTGAAGAGGATGCAGAAGTAACGACAATGGC  
2301 GCAATCAGTGCCATGGGGAATTAGCCGTGTGCAAGCCCCGGCTGCCATA  
2351 ACCGTGGATTGACAGGTTCTGGTGTAAAAGTTGCTGTCTCGATACAGGT

**SUBSTITUTE SHEET**



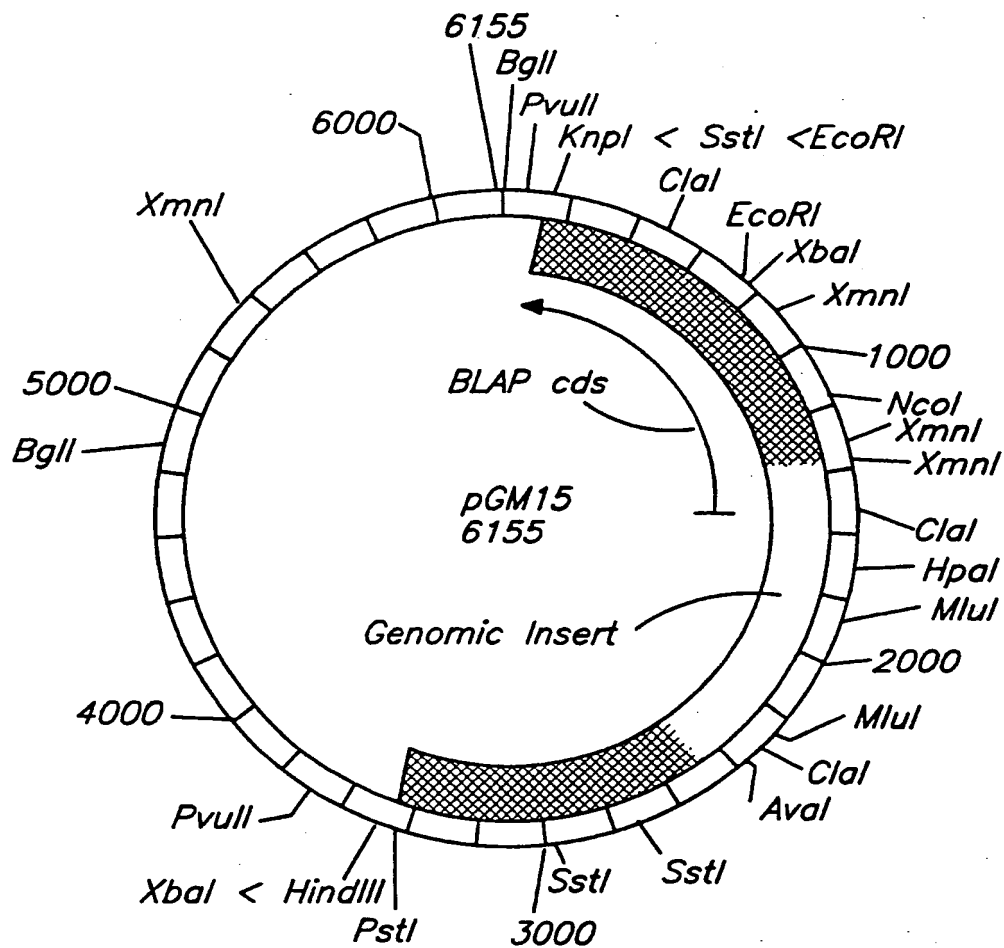
2401 ATTTCCACTCATCCAGACTTAAATATTTCGTGGTGGCGCTAGCTTTGTACC  
2451 AGGGGAACCATCCACTCAAGATGGGAATGGGCATGGCACGCATGTGGCCG  
2501 GGACGATTGCTGCTTTAAACAATTTCGATTGGCGTTCTTGGCGTAGCGCCT  
2551 AGTGCGGAACATACGCTGTTAAAGTTTTAGGAGCCGACGGTAGAGGTGC  
2601 AATCAGCTCGATTGCCCAAGGGTTGGAATGGGCAGGGAACAATGGCATGC  
2651 ACGTTGCTAATTTGAGTTTAGGAAGCCCTTCGCCAAGTGCCACACTTGAG  
2701 CAAGCTGTTAATAGCGCGACTTCTAGAGGCGTTCTTGTTGTAGCGGCATC  
2751 TGGGAATTCAGGTGCAAGCTCAATCAGCTATCCGGCCCCGTTATGCGAACG  
2801 CAATGGCAGTCGGAGCTACTGACCAAAACAACAACCGCGCCAGCTTTTCA  
2851 CAGTATGGCGCAGGGCTTGACATTGTGCGCACCAGGGGTAAACGTGCAGAG  
2901 CACATACCCAGGTTCAACGTATGCCAGCTTAAACGGTACATCGATGGCTA  
2951 CACCTCATGTTGCAGGTGCAGCAGCCCTTGTTAAACAAAAGAACCCTATCT  
3001 TGGTCCAATGTACAAATCCGCAACCATCTAAAGAATACGGCAACGAGCTT  
3051 AGGAAGCACGAACTTGTATGGAAGCGGACTTGTCAATGCAGAAGCGGCAA  
3101 CACGCTAATCAATAAAAAAACCGTGTGCGCTTAAAGGGCACAGCTTTTTT  
3151 TGTGTATGAATCGAAAAAAGAGAACAGATCGCAGGTCTCAAAAATCGAGC  
3201 GTAAAGGGTTGTTTAAAGCTCTTTACGCTCGCAGGTCTTATCGCTATACA  
3251 ATGGAAAATTCACGTCTTTTGACTTTTCATGGCATATATTTATTTAAGTAT  
3301 TCGTTTGCTTTTTTCGTACTCTCCGTTTTTCTGGTACCATTGCGCCAGCTC  
3351 AATTGCATAGTGGACTG6TTCTTCTTTATTATCAAGCTT

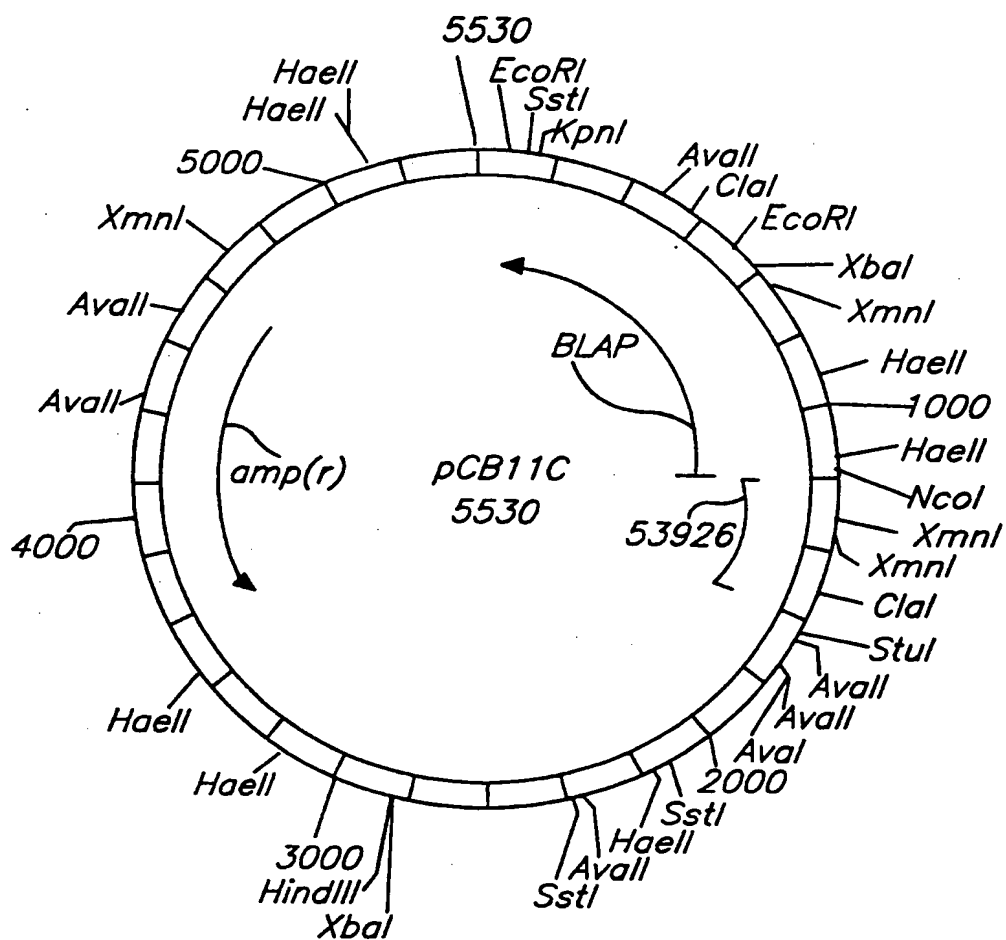
Total number of bases is: 3389.

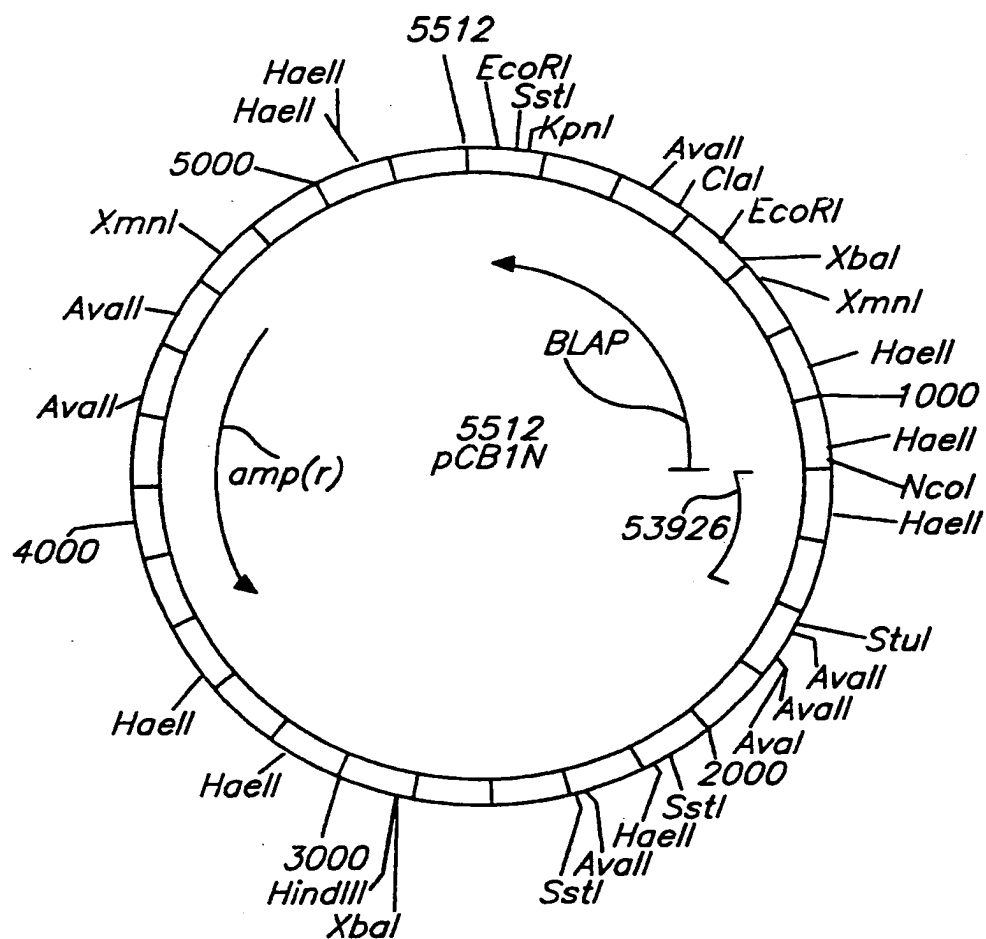
DNA sequence composition: 949 A; 716 C; 820 G; 904 T;

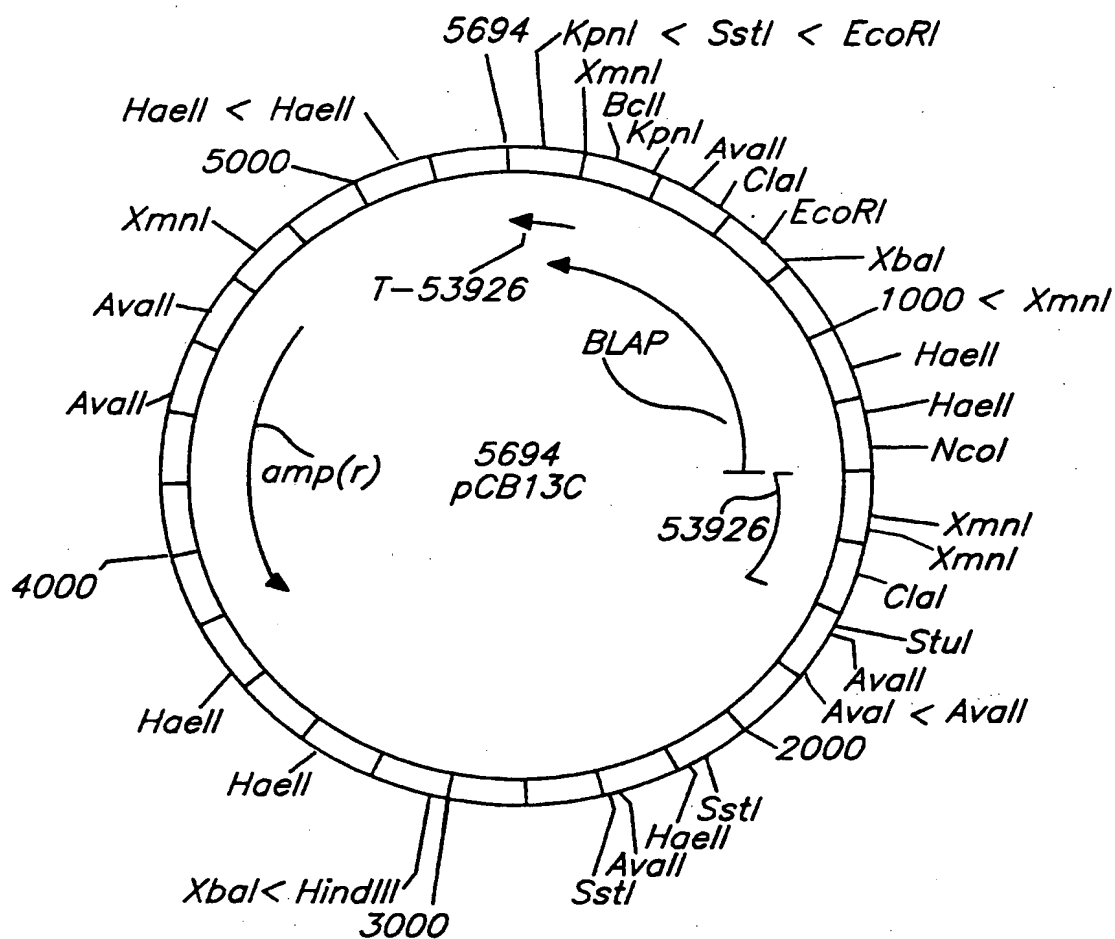
**FIG. 2-2**

**SUBSTITUTE SHEET**

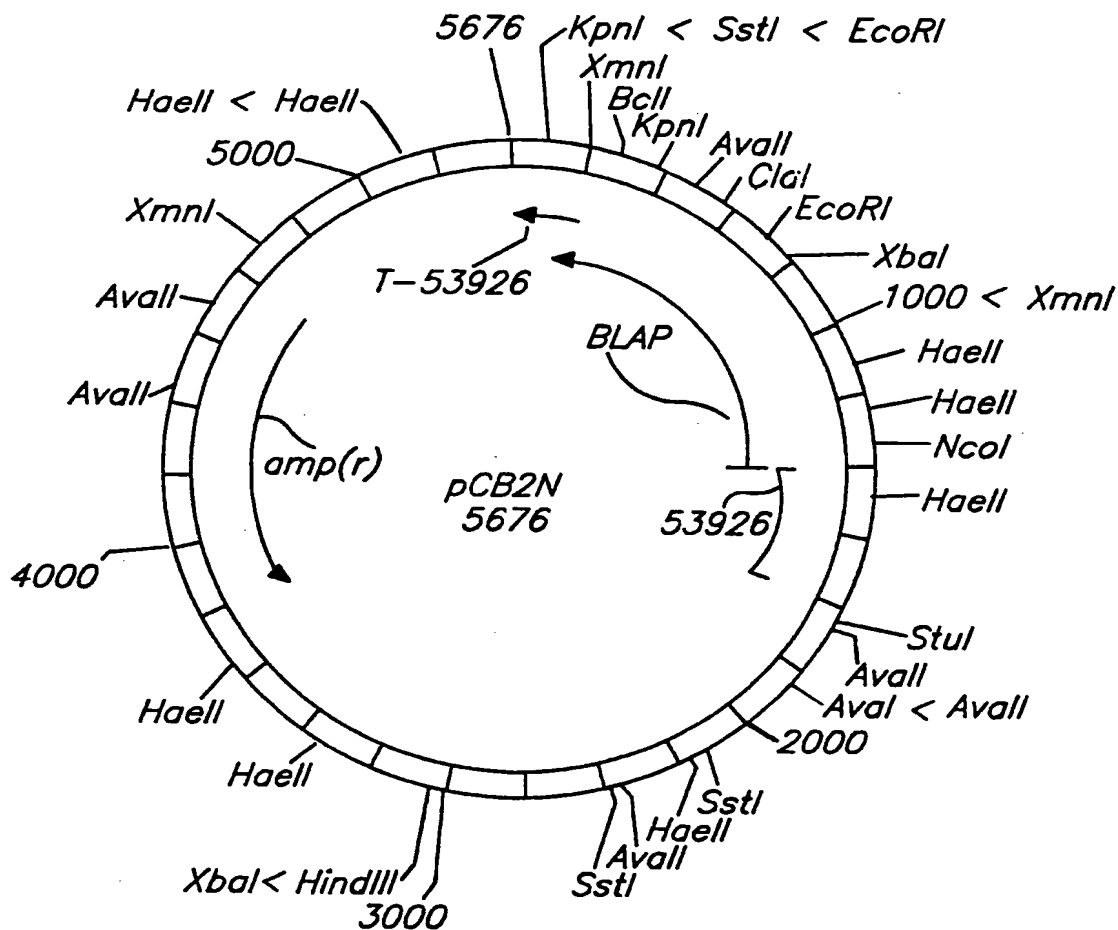
**FIG. 3****SUBSTITUTE SHEET**

5/31  
**FIG. 4****SUBSTITUTE SHEET**

**FIG. 5****SUBSTITUTE SHEET**

**FIG. 6**

SUBSTITUTE SHEET

**FIG. 7**

SUBSTITUTE SHEET

9/31

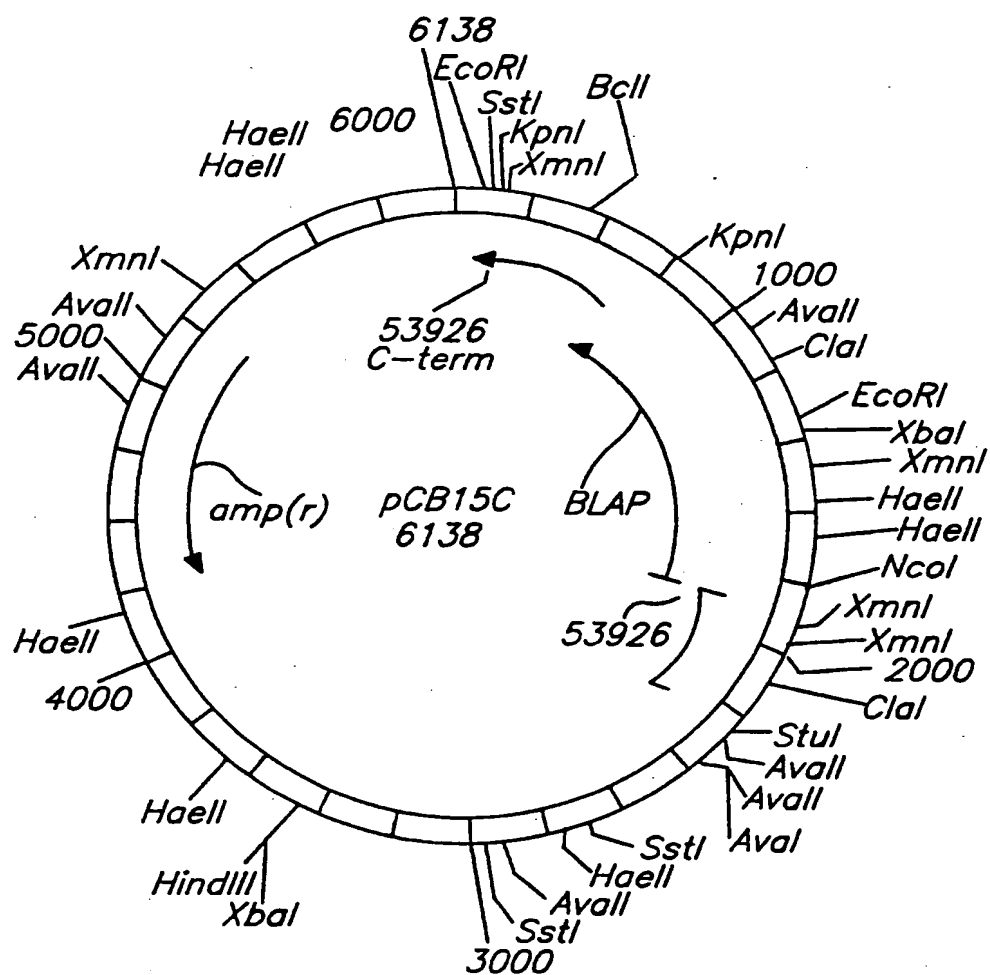
80 100  
#1 5' - GGCCCTCGGGACCTCTTTCCC - 3' (53926, top strand)  
Ava I

651 633  
#2 5' - GGAATGCCccATGGAACGG - 3' (53926, bottom strand)  
Nco I

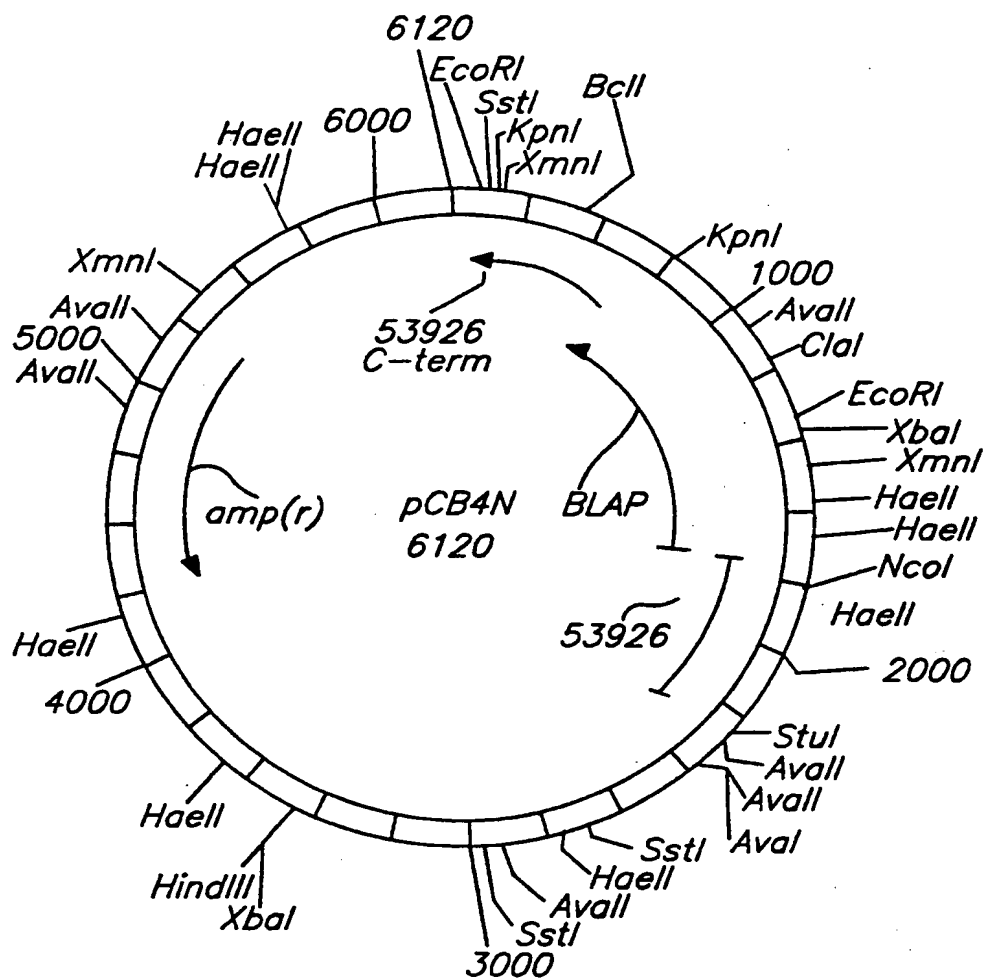
390 361  
#3 5' - GCGGAATCGaTcgATGCCATCGTGAACACG - 3' (53926, bottom strand)  
Cla I

**FIG. 8****SUBSTITUTE SHEET**

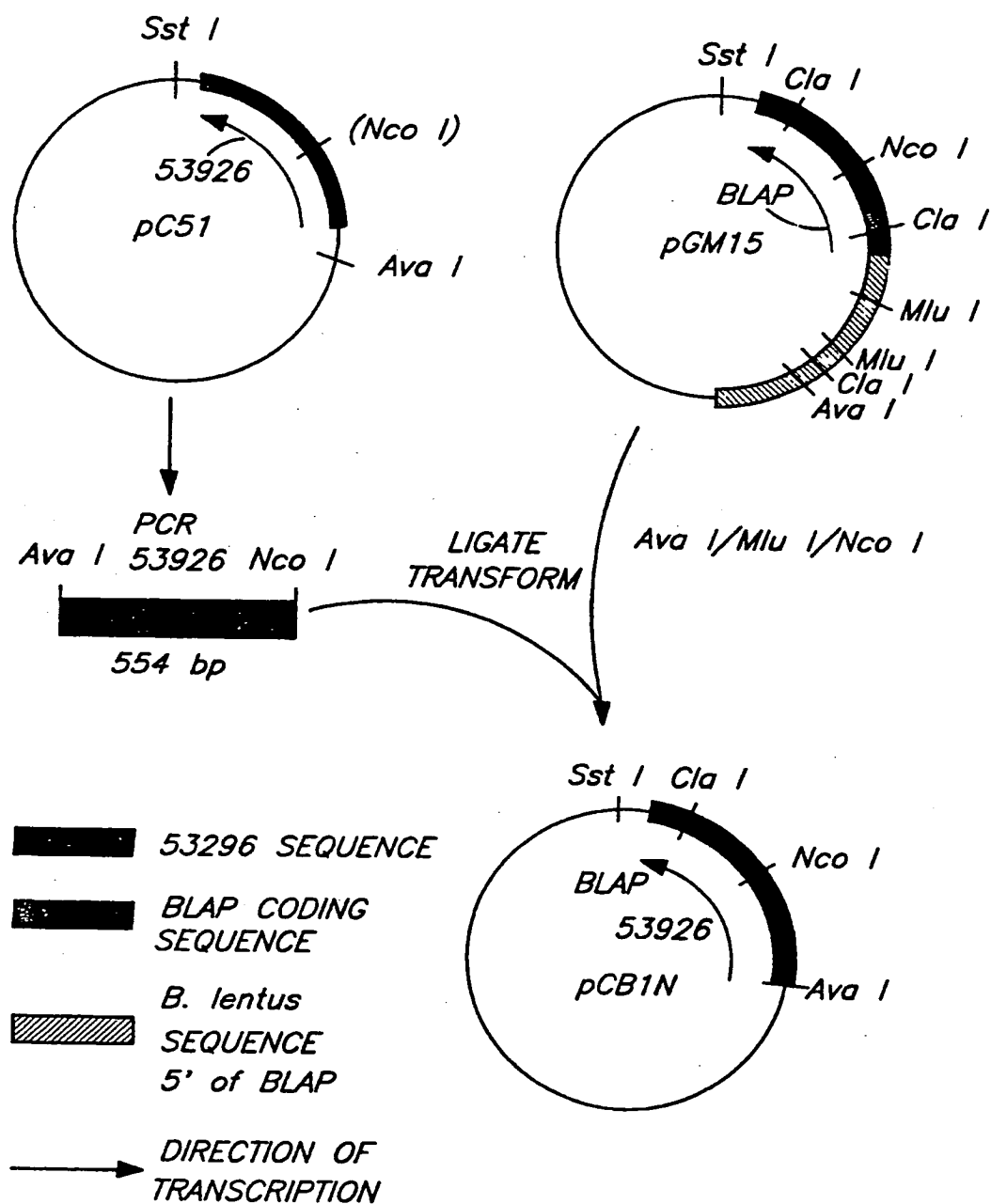
10/31

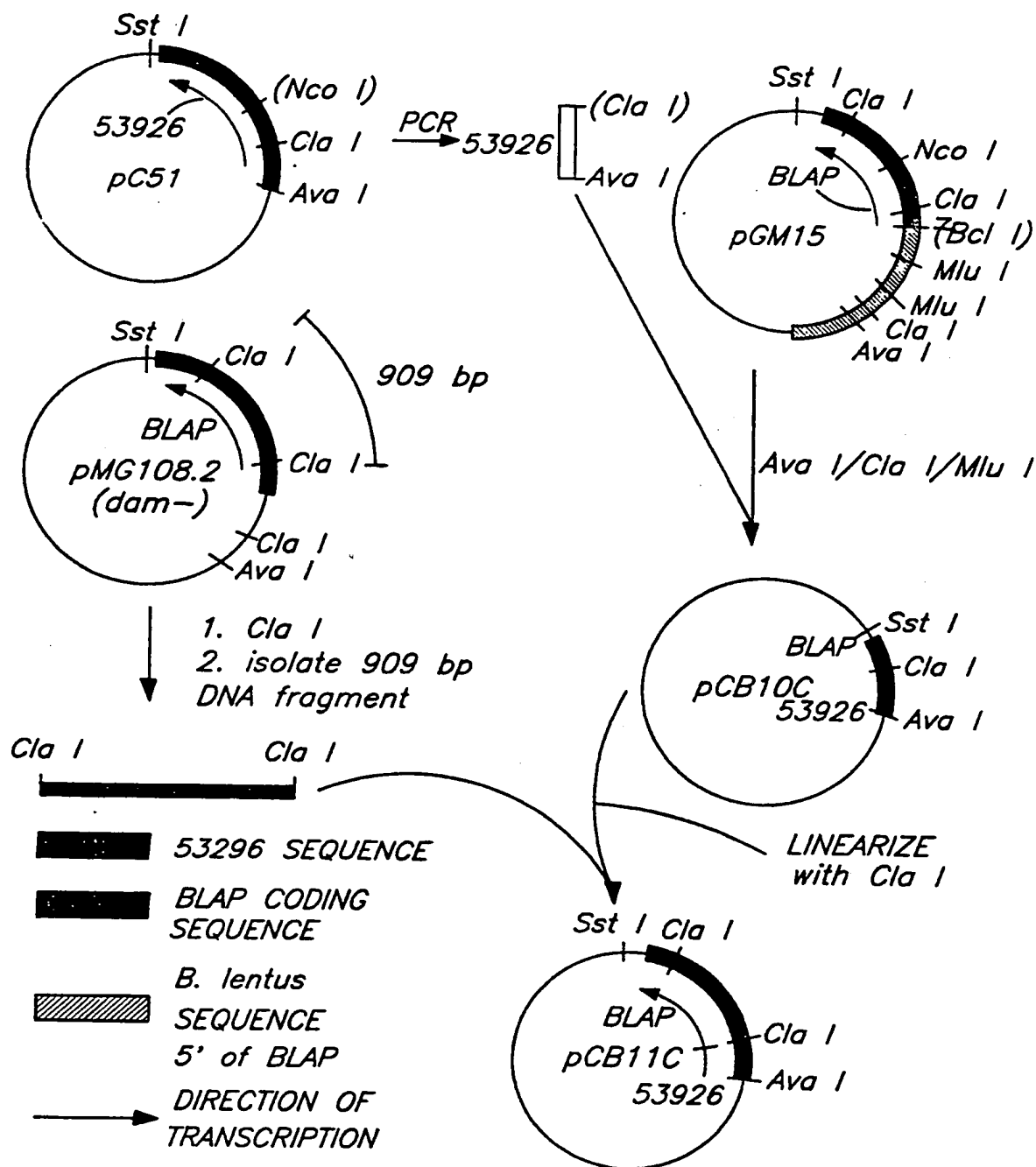
**FIG. 9****SUBSTITUTE SHEET**



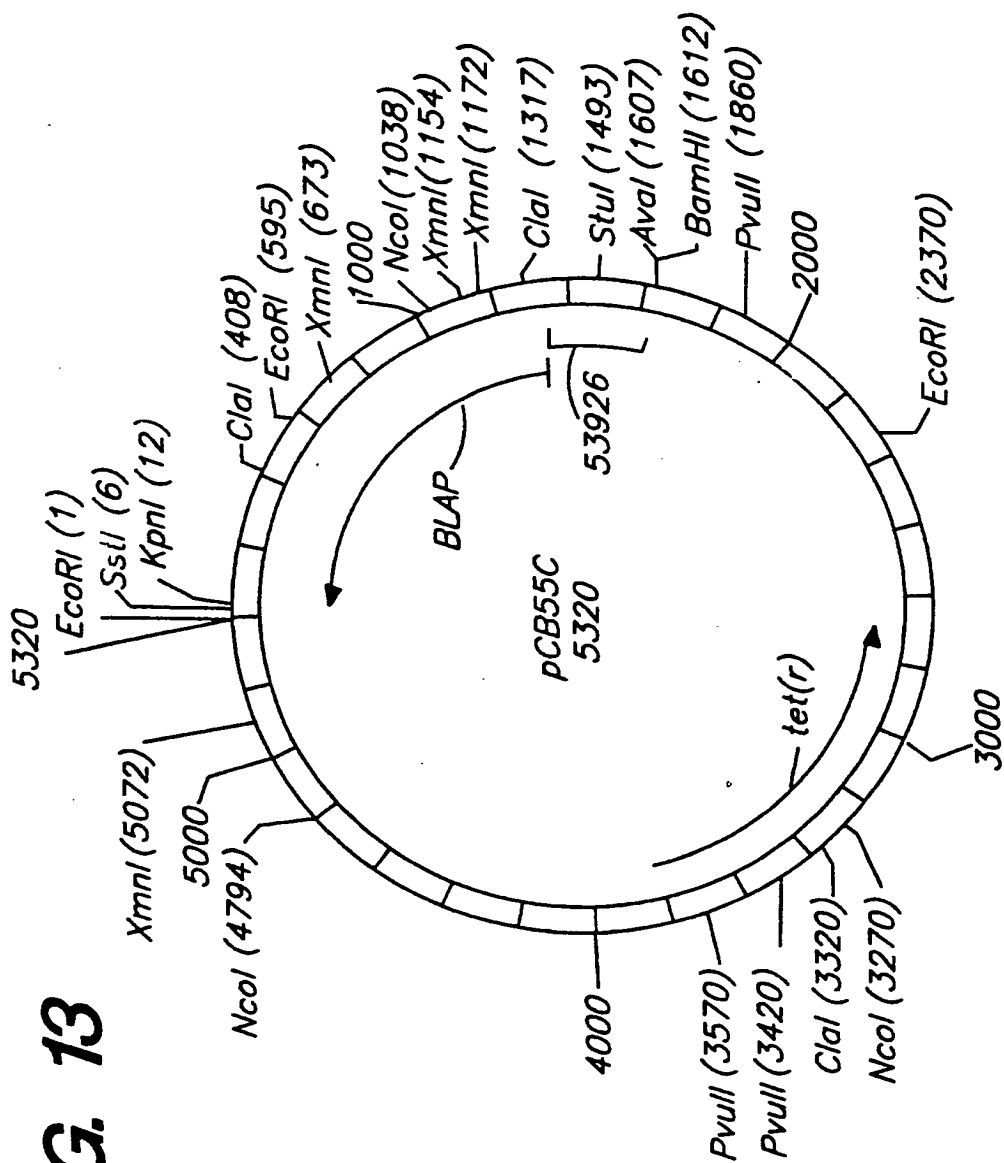
**FIG. 10****SUBSTITUTE SHEET**

12 / 31

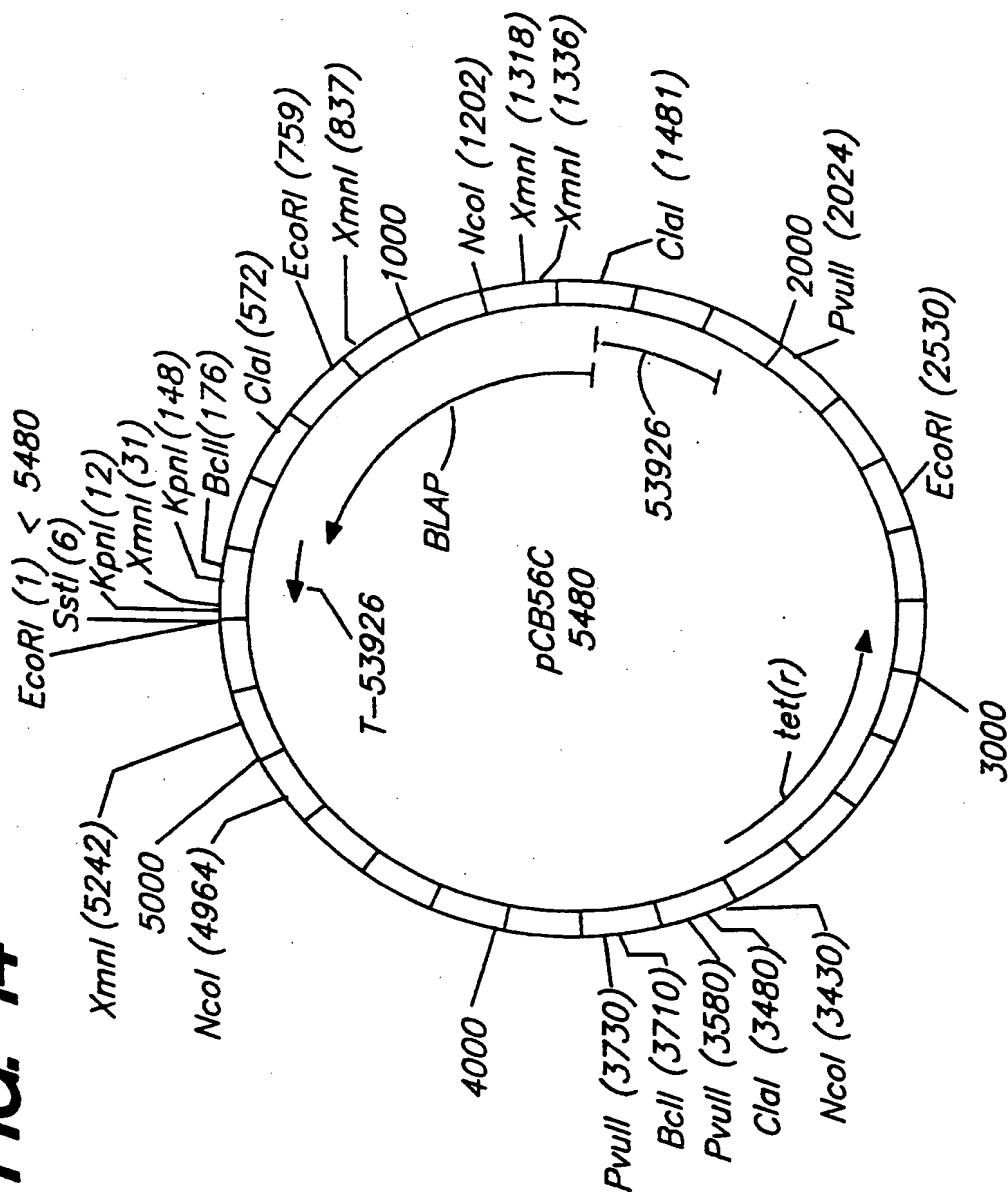
**FIG. 11****SUBSTITUTE SHEET**

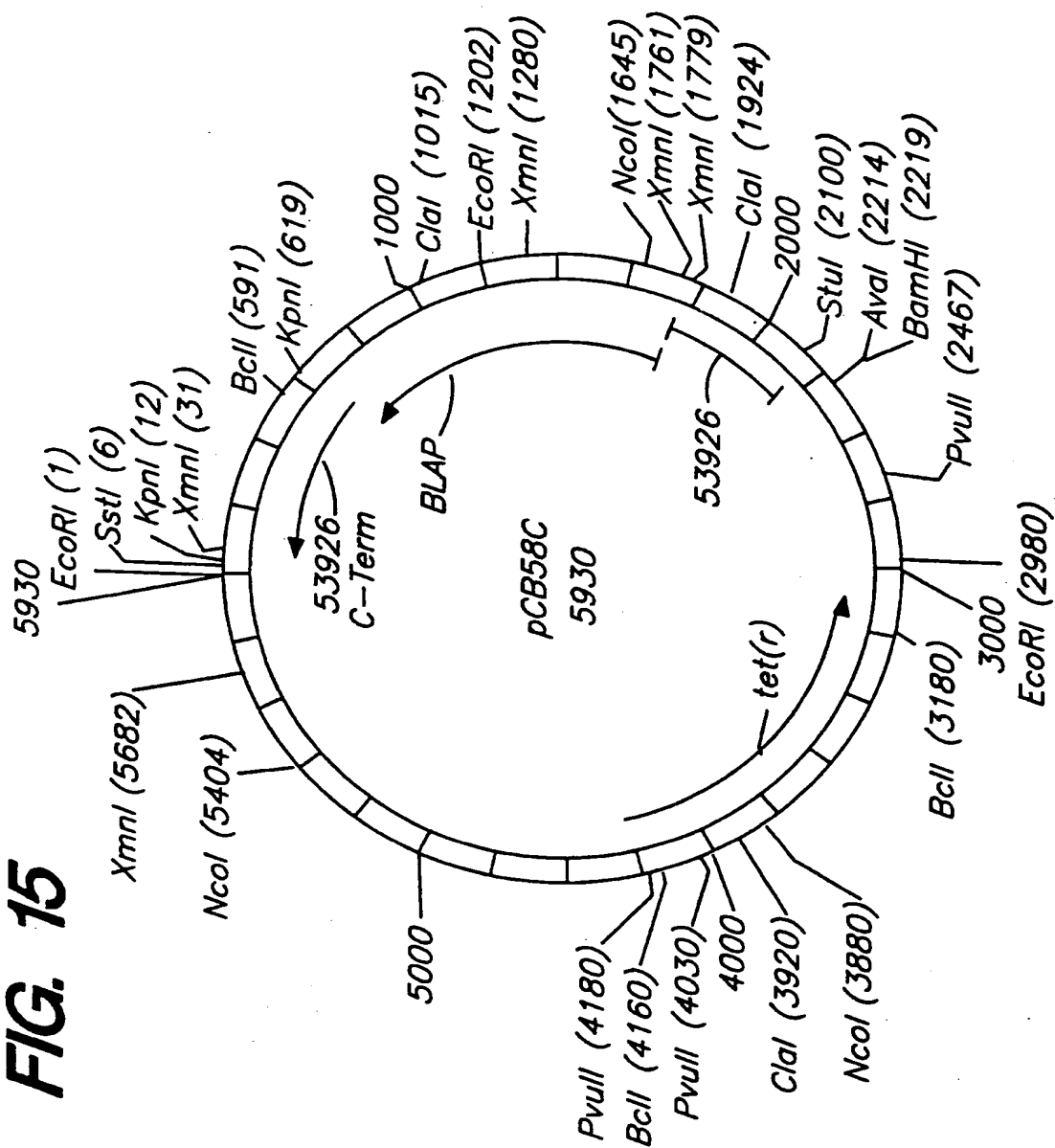
**FIG. 12**

SUBSTITUTE SHEET



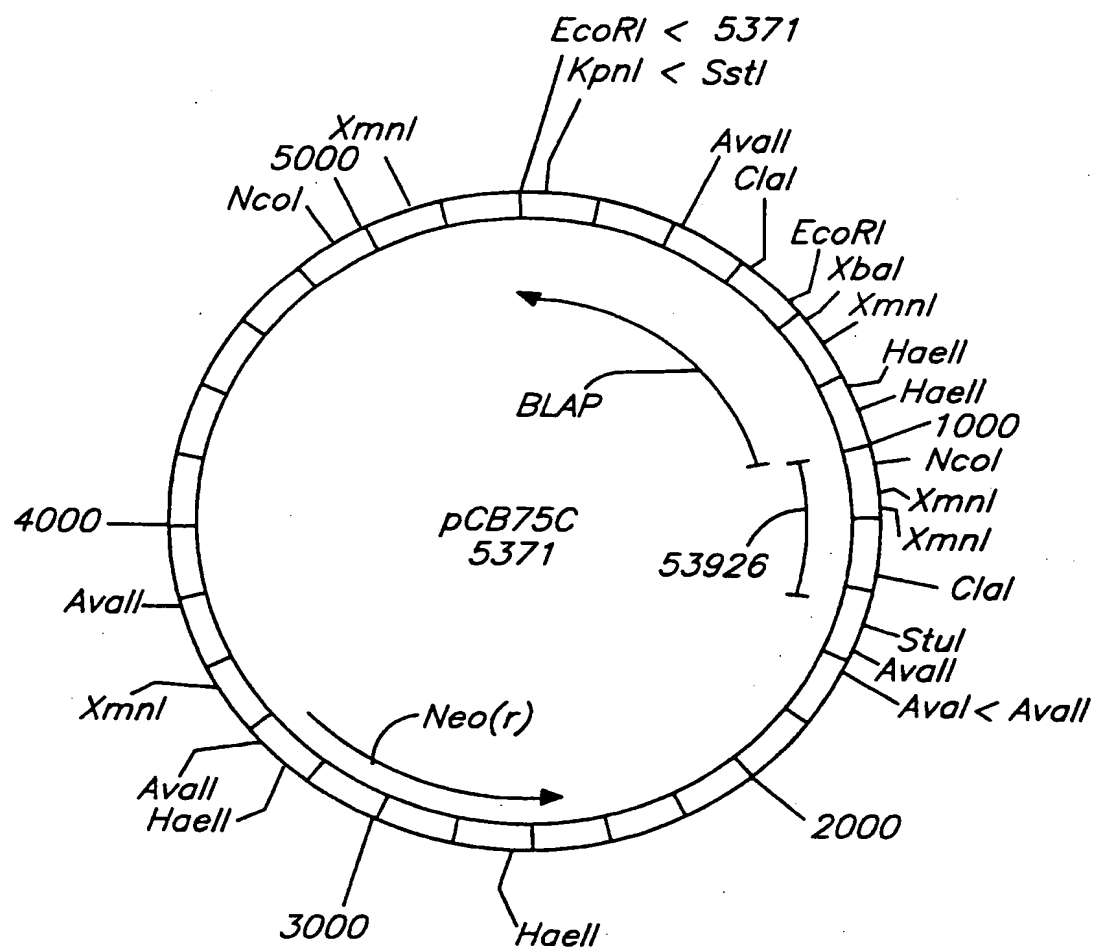
**FIG. 14**



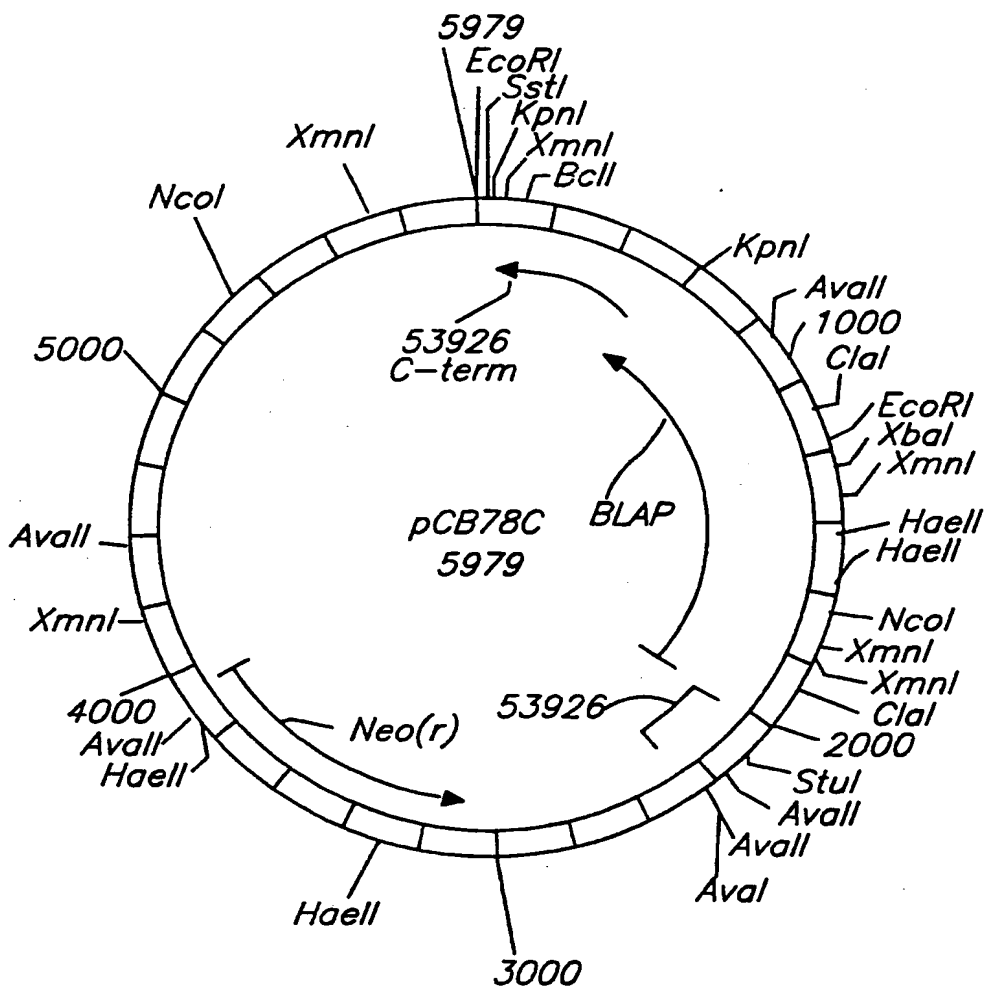
**FIG. 15**

SUBSTITUTE SHEET

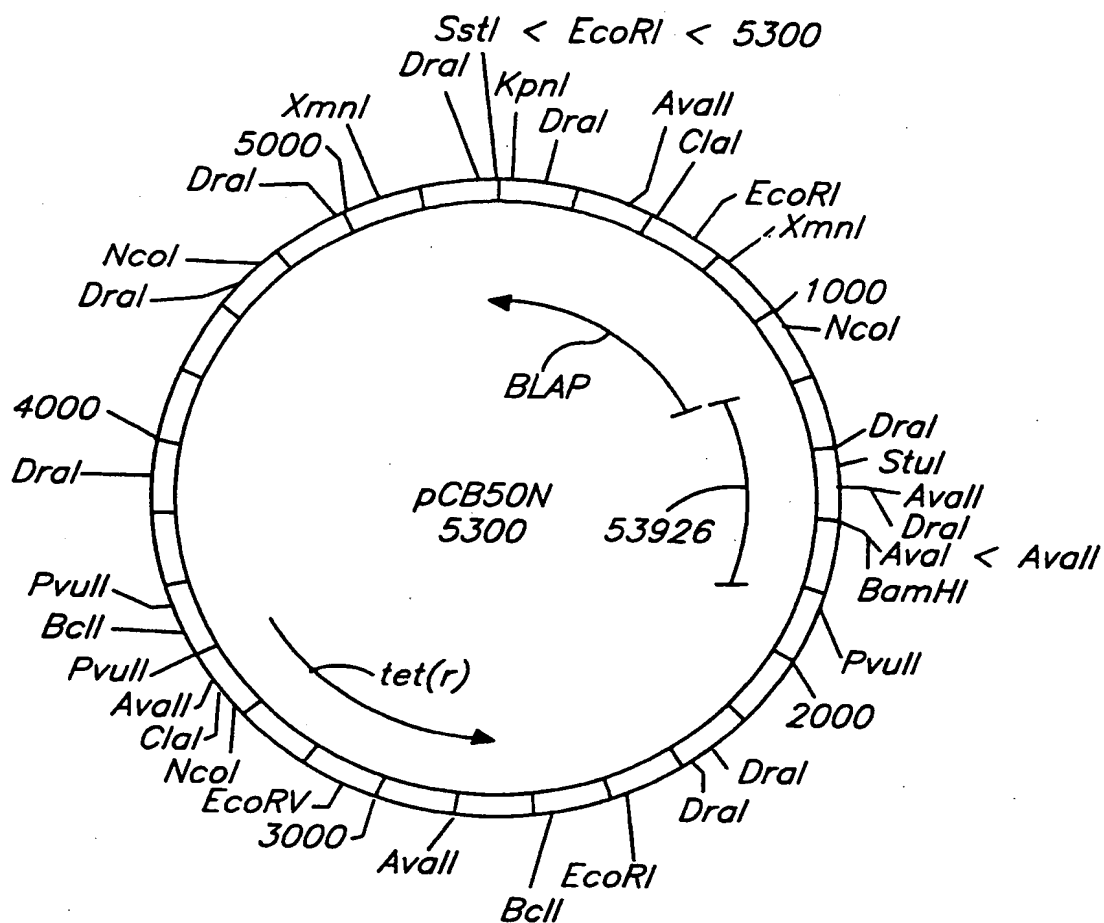
17/31

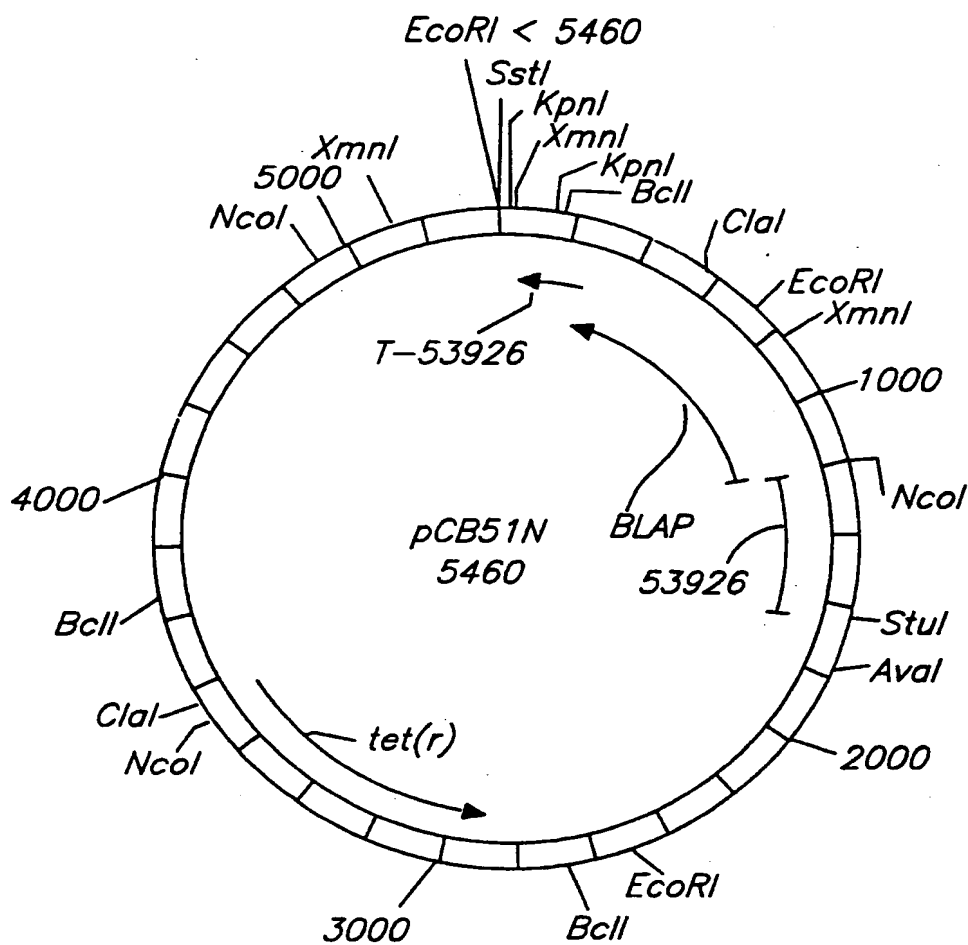
**FIG. 16****SUBSTITUTE SHEET**

**FIG. 17**

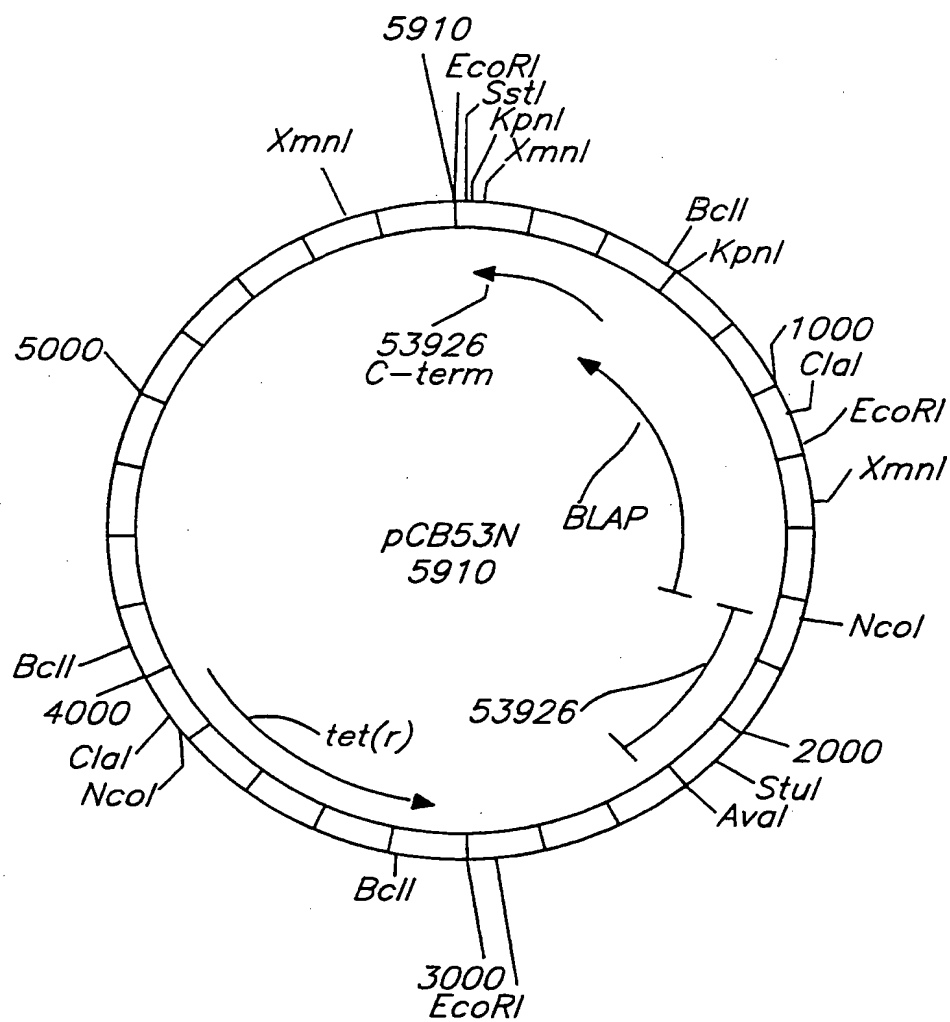


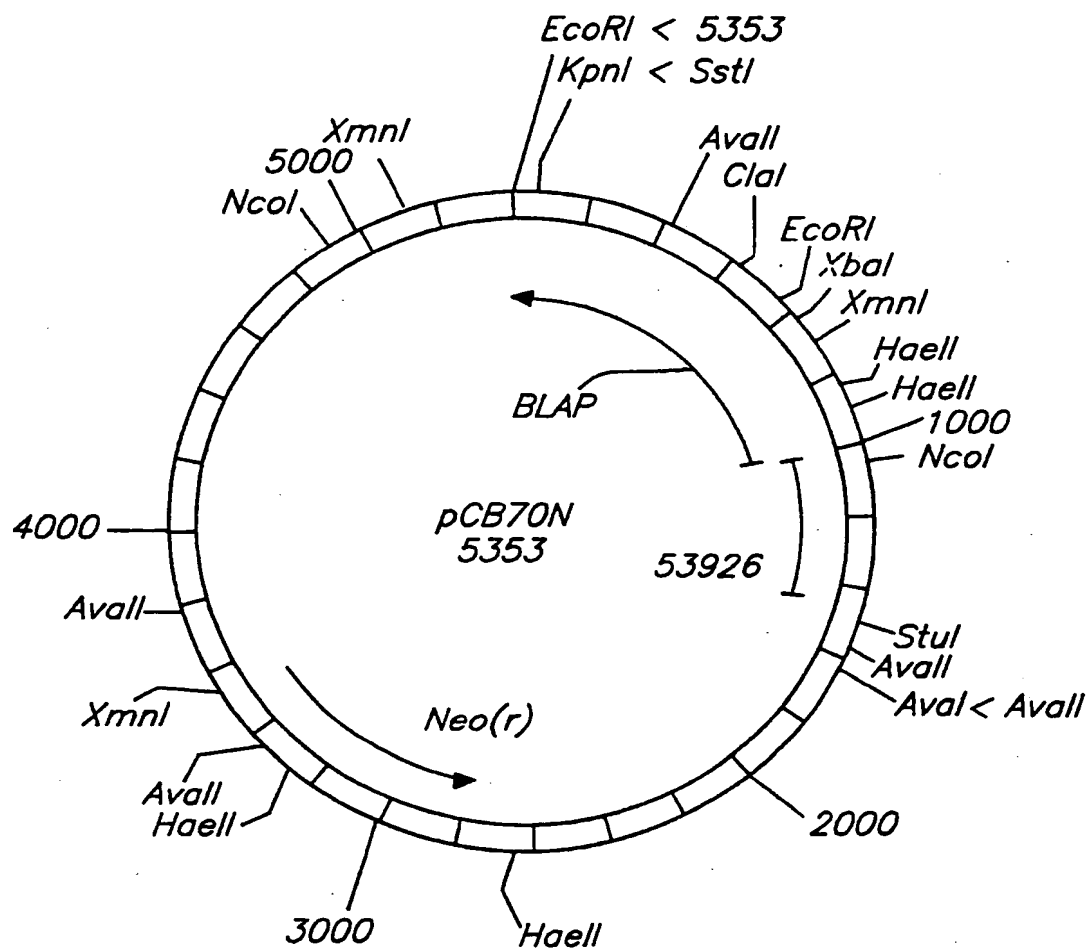


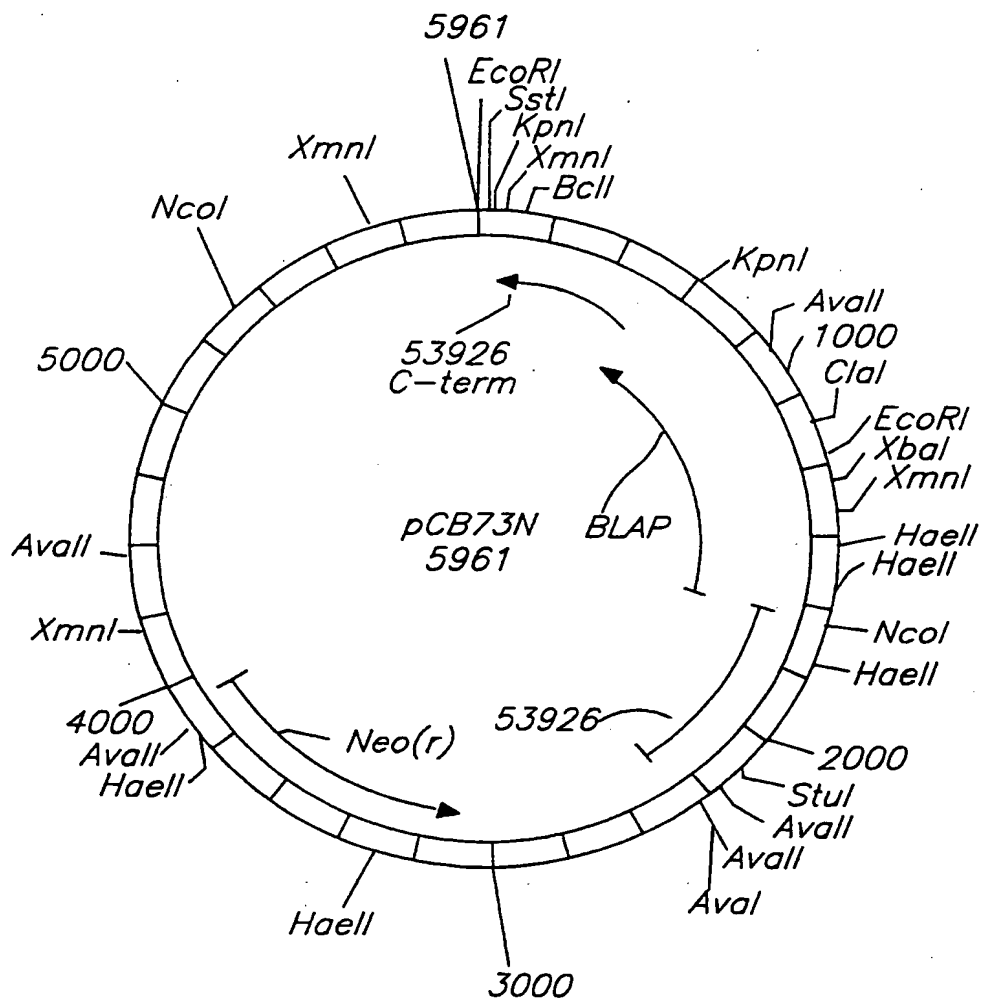
**FIG. 18****SUBSTITUTE SHEET**

**FIG. 19**

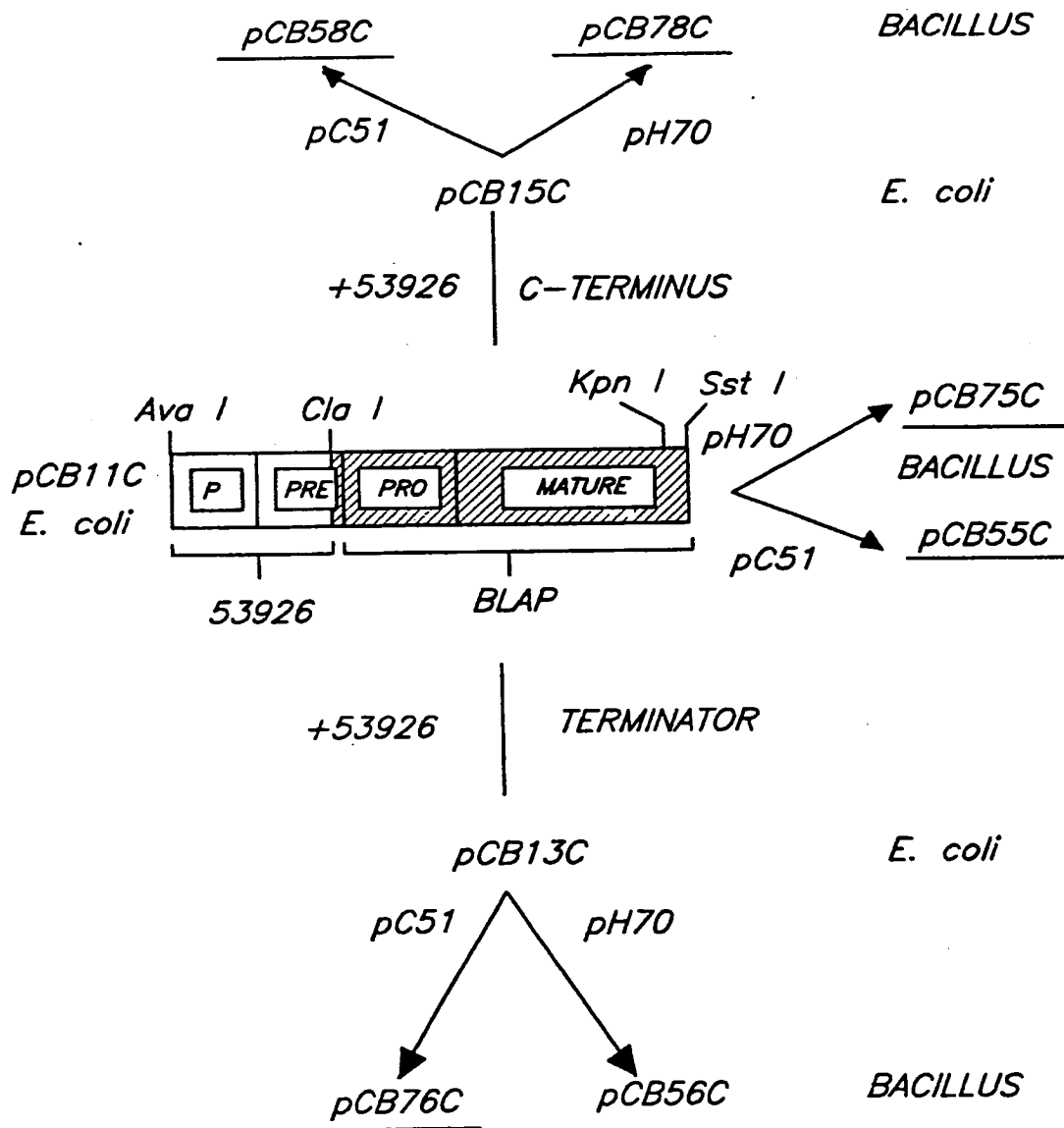
SUBSTITUTE SHEET

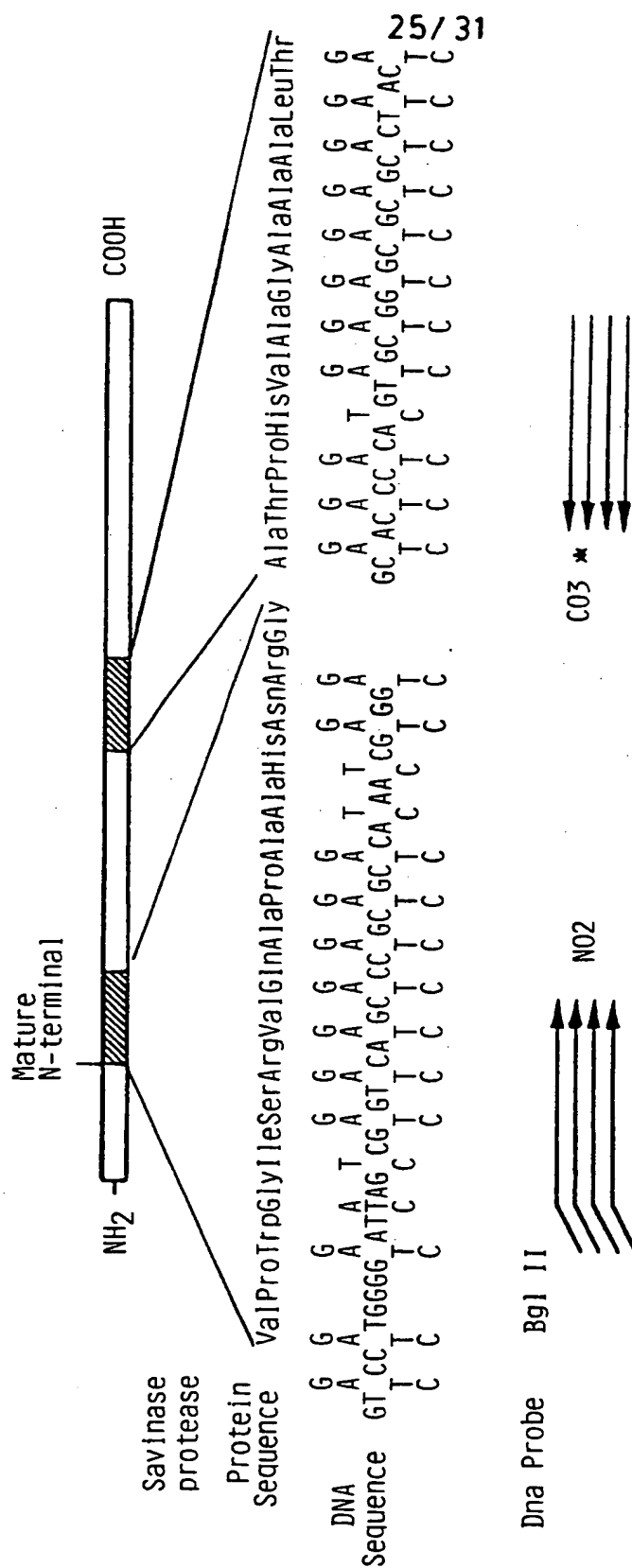
**FIG. 20****SUBSTITUTE SHEET**

**FIG. 21****SUBSTITUTE SHEET**

**FIG. 22****SUBSTITUTE SHEET**

**FIG. 23**





\*the sequence of the C03 series of probes is complementary to that shown

**FIG. 24**

# SUBSTITUTE SHEET

**FIG. 25**

N02.1 5' GGCAGATCTGTGCCXTGGGGGAT-3'  
N02.2 5' GGCAGATCTGTGCCXTGGGGAAT-3'  
N02.3 5' GGCAGATCTGTGCCXTGGGGTAT-3'  
N02.4 5' GGCAGATCTGTGCCXTGGGGCAT-3'

          A T       T  
N04.1 5' GTCCCGTGGGGXATCAG-3'

                  A  
          G T       T  
N04.2 5' GTTCCGTGGGGXATCAG-3'

                  A  
          A T       T  
N04.3 5' GTCCCGTGGGGXATCTC-3'

                  A  
          G T       T  
N04.4 5' GTTCCGTGGGGXATCTC-3'  
                  A

          A A A A T  
C03.1 5' CCCGCCACGTGCGGCGT-3'

          G G A A T  
C03.2 5' CCTGCTACGTGCGGCGT-3'

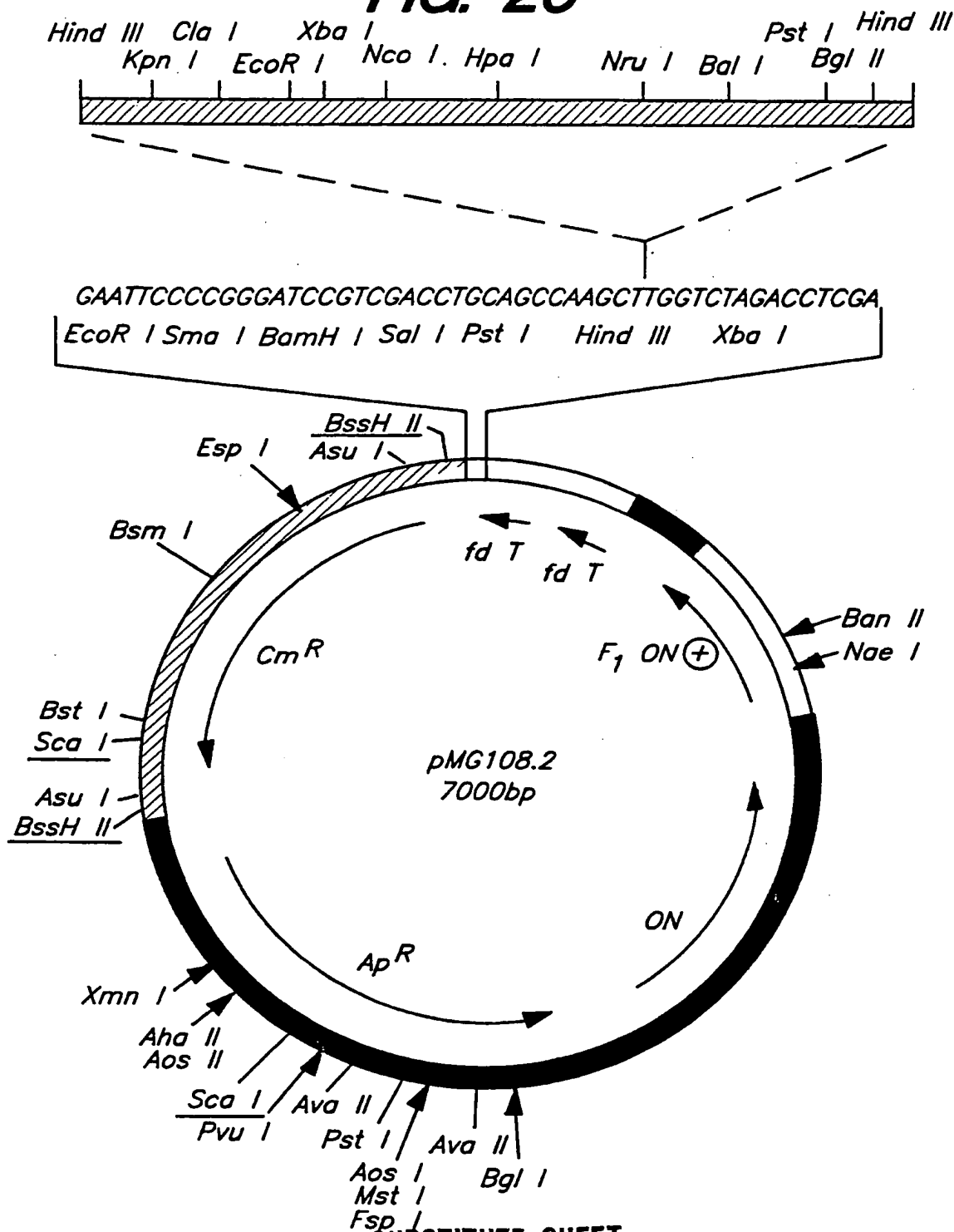
          G A A A T  
C03.3 5' CCTGCCACGTGCGGCGT-3'

          A G A A T  
C03.4 5' CCCGCTACGTGCGGCGT-3'

X = (A,T,G or C)

**SUBSTITUTE SHEET**



**FIG. 26**

**FIG. 27**

10 20 30 40 50

1 GTCTCCTGGATCTCATAAAATAAATGAATAGTATTTTCATAAAATGAATC

51 ATATGGATGCAATCTCCTGTCATTCTGCTGGCCCTCGGGACCTCTTTCCC

101 TCGGAGGCTGAAGCGGTCTATTCATACTTTTCTGAAGTGAACATTTTCTAA

151 AACAGTTATTAATAACCAAAAAATTTTAAATTGGTCCTCCAAAAAATAG

201 GCCTACCATATAATTCATTTTTTTTCTATAATAAATTAACAGATAATTG  
TRANSCRIPTION START SITE

251 GAATAGATTATATTATCCTTCTATTTAAATTATTCTGAATAAAGAGGAGG

301 AGAGTGAGTAATGATGAGGAAAAAGAGTTTTTGGCTTGGGATGCTGACGG  
PRE-SEQUENCE

351 CCTTCATGCTCGTGTTACGATGGCATTGCGATTCCGCTTCTGCTGCT  
CG T PRO SEQUENCE

401 CAACCGGCGAAAAATGTTGAAAAGGATTATATTGTCGGATTTAAGTCAGG  
Cla I

451 AGTGAAAACCGCATCTGTCAAAAAGGACATCATCAAAGAGAGCGGCGGAA

501 AAGTGGACAAGCAGTTTAGAATCATCAACGCGGCAAAAGCGAAGCTAGAC

551 AAAGAAGCGCTTAAGGAAGTCAAAAATGATCCGGATGTCGCTTATGTGGA  
MATURE PROTEASE

601 AGAGGATCATGTGGCCCATGCCTTGGCGCAAACCGTTCCCTTACGGCATTG  
A GG

651 CTCTCATTAAGCGGACAAAGTGCAGGCTCAAGGCTTTAAGGGAGCGAAT  
Nco I

701 GTAAAAGTAGCCGTCTCTGGATACAGGAATCCAAGCTTCTCATCCGGACTT

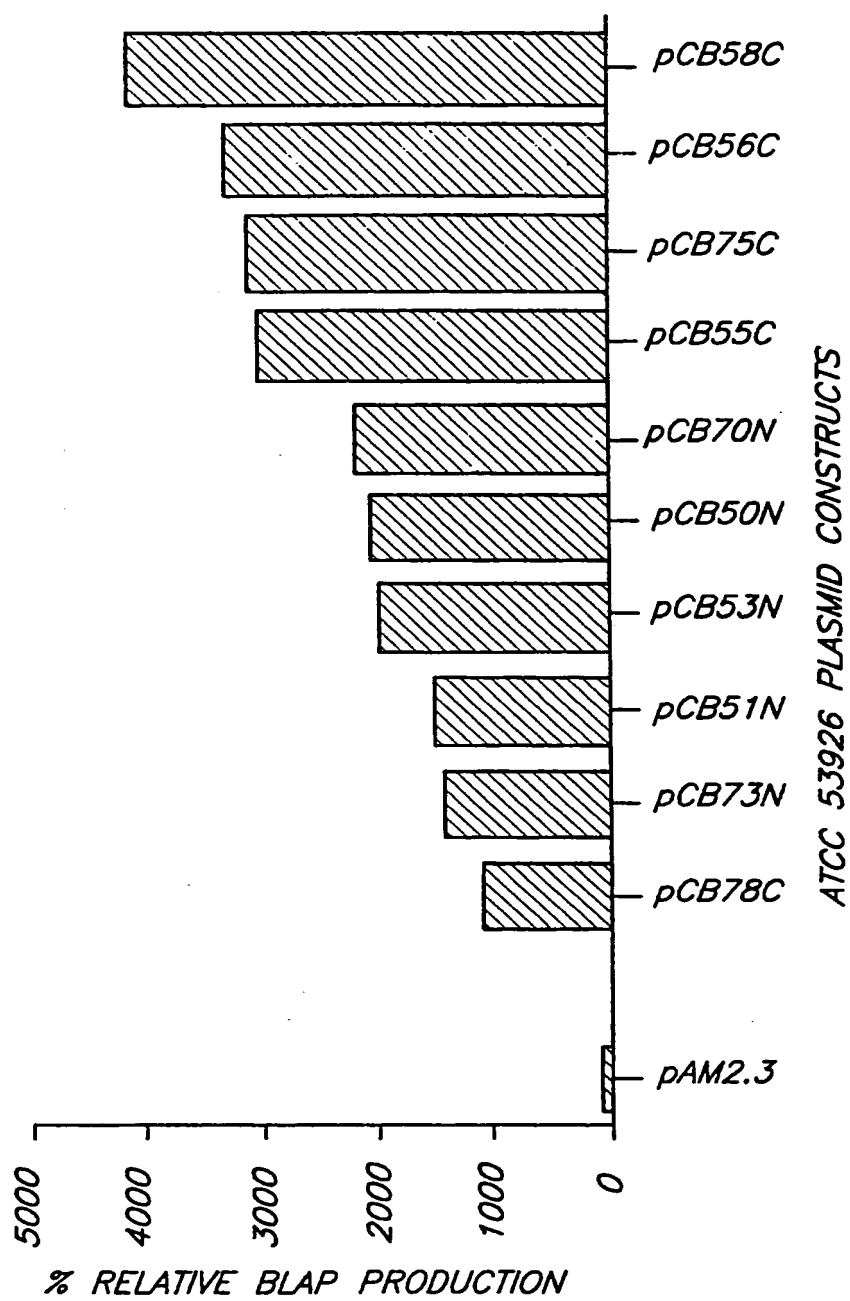
751 GAACGTAGTCGGCGGAGCAAGCTTTGTGGCTGGCGAAGCTTATAACACCG

**SUBSTITUTE SHEET**

29/ 31

801 ACGGCAACGGACACGGCACACATGTTGCCGGTACAGTAGCTGCGCTTGAC  
851 AATACAACGGGTGTATTAGGCGTTGCGCCAAGCGTATCCTTGACGCGGT  
901 TAAAGTACTGAATTCAAGCGGAAGCGGATCATACAGCGGCATTGTAAGCG  
T C  
951 GAATCGAGTGGGCGACAACAAACGGCATGGATGTTATCAATATGAGCCTT  
Kpn I  
1001 GGGGAGCATCAGGCTCGACAGCGATGAAACAGGCAGTCGACAATGCATA  
1051 TGCAAGAGGGGTTGTCGTTGTAGCTGCAGCAGGGAACAGCGGATCTTCAG  
1101 GAAACACGAATACAATTGGCTATCCTGCAAAATACGATTCTGTCATCGCT  
1151 GTTGGTGCGGTAGACTCTAACAGCAACAGAGCTTCATTTTCCAGCGTCGG  
1201 AGCAGAGCTTGAAGTCATGGCTCCTGGCGCAGGCGTATACAGCACTTACC  
1251 CAACGAACACTTATGCAACATTGAACGGAACGTCAATGGCTTCTCCTCAT  
1301 GTAGCGGGAGCAGCAGCTTTGATCTTGTCAAACATCCGAACCTTTCAGC  
T C  
1351 TTCACAAGTCCGCAACCGTCTCTCCAGCACGGCGACTTATTTGGGAAGCT  
Kpn I  
1401 CCTTCTACTATGGGAAAGGTCTGATCAATGTCGAAGCTGCCGCTCAATAA  
1451 CATATTCTAACAAATAGCATATAGAAAAAGCTAGTGTTTTTAGCACTAGC  
1501 TTTTTCTTCATTCTGATGAAGGTTGTTCAATATTTTGAATCCGTTCCATG  
G AC  
1551 ATCGTCGGATGGCCGTATTTAAAAATCTTGACGAGAAACGGCGGGTTTGC  
Kpn I  
1601 CTCGCTCAGCCCGCTTTTGAGAGCTCTTGAAACGTCGAAACCGCTGCAT  
1651 CGCTGTTTTGCGTCAGTTCAATCGCATACTGGTCAGCAGCTTTTCCTGAT  
1701 GCCTGAAACTGCTGCTAAATGAGACGA

**FIG. 27A****SUBSTITUTE SHEET**

**FIG. 28**

SUBSTITUTE SHEET

		5		10		15		20		25		30																		
1	A	Q	T	V	P	W	G	I	S	R	V	Q	A	P	A	A	H	N	R	G	L	T	G	S	G	V	K	V	A	V
31	L	D	T	G	I	S	T	H	P	D	L	N	I	R	G	G	A	S	F	V	P	G	E	P	S	T	Q	D	G	N
61	G	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	E	L	Y	A
91	V	K	V	L	G	A	D	G	R	G	A	I	S	S	I	A	Q	G	L	E	W	A	G	N	N	G	M	H	V	A
121	N	L	S	L	G	S	P	S	P	S	A	T	L	E	Q	A	V	N	S	A	T	S	R	G	V	L	V	V	A	A
151	S	G	N	S	G	A	S	S	I	S	Y	P	A	R	Y	A	N	A	M	A	V	G	A	T	D	Q	N	N	N	R
181	A	S	F	S	Q	Y	G	A	G	L	D	I	V	A	P	G	V	N	V	Q	S	T	Y	P	G	S	T	Y	A	S
211	L	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	V	K	Q	K	N	P	S	W	S	N	V	Q	I	
241	R	N	H	L	K	N	T	A	T	S	L	G	S	T	N	L	Y	G	S	G	L	V	N	A	E	A	A	T	R	

## Amino acid composition

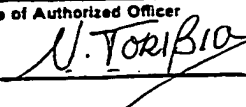
41 A	0 C	7 H	3 M	18 T
9 R	10 Q	10 I	2 F	3 W
22 N	5 E	19 L	13 P	7 Y
6 D	34 G	5 K	31 S	24 V

Number of residues: 269  
Molecular weight (MW): 26837

**FIG. 29****SUBSTITUTE SHEET**

# INTERNATIONAL SEARCH REPORT

International Application No PCT/US 90/04673

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (if several classification symbols apply, indicate all) <sup>6</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC IPC <sup>5</sup> : C 12 N 9/54, 15/57, 15/75, C 11 D 3/386, //(C 12 N 1/21, C 12 R 1:10)(C 12 N 1/20, C 12 R 1:07, 1:10)		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>7</sup>		
Classification System	Classification Symbols	
IPC <sup>5</sup>	C 12 N	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched <sup>8</sup>		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT <sup>9</sup></b>		
Category <sup>10</sup>	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
Y	WO, A, 89/06279 (NOVO INDUSTRI A/S) 13 July 1989 see abstract; table 1a,; page 14, line 20 - page 17, line 24	1,36,47
A	--	2-5,16-25
Y	EP, A, 0214435 (HENKEL KOMMANDITGESELL- SCHAFT AUF AKTIEN) 18 March 1987 see abstract; page 2, lines 10-15; all claims	1,36,47
A	--	2-4,16-25
A	GB, A, 2171703 (AGENCY OF INDUSTRIAL SCIENCE AND TECHNOLOGY) 3 September 1986 see abstract	2-4
-----		
<p><sup>10</sup> Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"A" document member of the same patent family</p>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report	
26th Oktober 1990	17. 12. 90	
International Searching Authority	Signature of Authorized Officer	
EUROPEAN PATENT OFFICE	 Nuria TORIBIO	

**ANNEX TO THE INTERNATIONAL SEARCH REPORT  
ON INTERNATIONAL PATENT APPLICATION NO.**

US 9004673  
SA 39530

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 05/12/90. The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A- 8906279	13-07-89	EP-A- 0396608	14-11-90
EP-A- 0214435	18-03-87	DE-A- 3527913	12-02-87
		AU-A- 6078986	05-02-87
		JP-A- 62032888	12-02-87
GB-A- 2171703	03-09-86	JP-A- 61166400	28-07-86
		JP-A- 61212288	20-09-86

EPO FORM P0079

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

**THIS PAGE BLANK (USPTO)**